GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 3, 2003, 09:54:17 ; Search time 304 Seconds
(without alignments)
1340.366 Million cell updates/sec

Title: US-09-881-736-2
Perfect score: 3243
Sequence: 1 MDTMMLNVRNLFEQLVRRVE.....SKSATNLGRQGNFFASPMLK 632

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3243	100.0	632	21	US-09-791-537-1796	Sequence 1796, Ap
2	3243	100.0	632	22	US-09-881-736-2	Sequence 2, Appli
3	3237	99.8	632	20	US-09-641-377-585	Sequence 585, App
4	3233	99.7	632	18	US-09-488-725A-3192	Sequence 3192, Ap
5	3230	99.6	632	21	US-09-791-537-72082	Sequence 72082, A
6	3230	99.6	632	22	US-09-833-790-413	Sequence 413, App
7	2919	90.0	570	22	US-09-833-790-427	Sequence 427, App
8	2734	84.3	628	21	US-09-791-537-36304	Sequence 36304, A
9	2734	84.3	628	22	US-09-881-736-4	Sequence 4, Appli
10	1182	36.4	255	20	US-09-629-469A-14587	Sequence 14587, A
11	1182	36.4	255	27	US-60-389-987-1045	Sequence 1045, Ap
12	1182	36.4	255	27	US-60-412-418-1045	Sequence 1045, Ap
13	1103.5	34.0	243	21	US-09-758-472-6751	Sequence 6751, Ap
14	1103.5	34.0	243	26	US-10-235-926-6751	Sequence 6751, Ap
15	971.5	30.0	256	1	PCT-US01-08656-9072	Sequence 9072, Ap
16	940.5	29.0	201	27	US-60-182-569-1026	Sequence 1026, Ap
17	889.5	27.4	191	27	US-60-182-569-1025	Sequence 1025, Ap
18	864	26.6	625	20	US-09-614-150-30960	Sequence 30960, A
19	864	26.6	625	21	US-09-791-537-49124	Sequence 49124, A
20	864	26.6	625	27	US-60-191-637-30526	Sequence 30526, A
21	749.5	23.1	165	27	US-60-186-662-967	Sequence 967, App
22	662.5	20.4	148	27	US-60-186-662-968	Sequence 968, App
23	646	19.9	187	27	US-60-207-215-294	Sequence 294, App
24	632	19.5	142	27	US-60-186-655-696	Sequence 696, App
25	631	19.5	872	1	PCT-US01-08656-9071	Sequence 9071, Ap
26	631	19.5	872	27	US-60-311-261-2570	Sequence 2570, Ap
27	625	19.3	4318	1	PCT-US01-08631-45146	Sequence 45146, A
28	618	19.1	665	1	PCT-US01-08631-40589	Sequence 40589, A
29	618	19.1	1086	1	PCT-US01-08631-50263	Sequence 50263, A
30	618	19.1	1139	1	PCT-US01-08631-44317	Sequence 44317, A
31	581	17.9	119	18	US-09-488-725A-6764	Sequence 6764, Ap
32	549	16.9	135	27	US-60-162-245-3866	Sequence 3866, Ap
33	549	16.9	135	27	US-60-169-868-6453	Sequence 6453, Ap
34	547.5	16.9	681	21	US-09-791-537-76252	Sequence 76252, A
35	541.5	16.7	681	22	US-09-881-736-6	Sequence 6, Appli
36	521	16.1	114	27	US-60-162-245-3547	Sequence 3547, Ap
37	477	14.7	383	21	US-09-791-537-35867	Sequence 35867, A
38	477	14.7	383	21	US-09-791-537-77334	Sequence 77334, A
39	477	14.7	384	20	US-09-614-150-4533	Sequence 4533, Ap
40	477	14.7	384	21	US-09-791-537-65301	Sequence 65301, A
41	477	14.7	384	27	US-60-167-217-4632	Sequence 4632, Ap
42	477	14.7	384	27	US-60-191-637-4551	Sequence 4551, Ap
43	477	14.7	385	21	US-09-791-537-150251	Sequence 150251,
44	460	14.2	367	27	US-60-146-315-901	Sequence 901, App
45	334	10.3	2022	27	US-60-389-987-598	Sequence 598, App

ALIGNMENTS

RESULT 1
US-09-791-537-1796
; Sequence 1796, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1796
; LENGTH: 632
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-1796

Query Match 100.0%; Score 3243; DB 21; Length 632;
Best Local Similarity 100.0%; Pred. No. 2.8e-263;
Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60

Qy         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120
```

```
Db         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120

Qy        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180

Db        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180

Qy        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240

Db        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240

Qy        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300

Db        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300

Qy        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360

Db        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360

Qy        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420

Db        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420

Qy        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480

Db        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480

Qy        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540

Db        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540

Qy        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600

Db        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600

Qy        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632

Db        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
```

RESULT 2
US-09-881-736-2
; Sequence 2, Application US/09881736
; GENERAL INFORMATION:
  APPLICANT: Glotzer, Michael
  APPLICANT: Jantsch-Plunger, Verena
  APPLICANT: Romano, Alper
  APPLICANT: Mishima, Masanori
; APPLICANT: Kaitna, Susanne
; TITLE OF INVENTION: Cyk-4 polypeptides, DNA molecules encoding them and their use in
; TITLE OF INVENTION: screening methods
; FILE REFERENCE: 0652.2260001/EKS/AES
; CURRENT APPLICATION NUMBER: US/09/881,736
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: EP 00 112 880.0
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: EP 01 110 554.1
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/241,231
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: To be determined
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 632
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-881,736-2

Query Match 100.0%; Score 3243; DB 22; Length 632;
Best Local Similarity 100.0%; Pred. No. 2.8e-263;
Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60

Db          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60

Qy         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120

Db         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120

Qy        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180

Db        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180

Qy        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240

Db        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240

Qy        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300

Db        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300

Qy        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360

Db        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360

Qy        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420

Db        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420

Qy        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480

Db        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480

Qy        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540

Db        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540

Qy        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600

Db        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600

Qy        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632

Db        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
```

RESULT 3
US-09-641-377-585
; Sequence 585, Application US/09641377
; GENERAL INFORMATION:
; APPLICANT: WIEMANN, STEFAN
; APPLICANT: GASSENHUBER, JOHANN
; APPLICANT: TAMPE, JENS
; TITLE OF INVENTION: HUMAN DNA SEQUENCES
; FILE REFERENCE: 087100/0106
; CURRENT APPLICATION NUMBER: US/09/641,377
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,499
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: 60/156,503
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 1793
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 585
; LENGTH: 632
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-641-377-585

Query Match 99.8%; Score 3237; DB 20; Length 632;
Best Local Similarity 99.8%; Pred. No. 8.9e-263;
Matches 631; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60

Qy      61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120

Qy     121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180

Qy     181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240

Qy     241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300

Qy     301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360

Qy     361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420

Qy     421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480

Qy     481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMLQDIKRQPKVVERLLSLPLEYWS 540

Qy     541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600

Qy     601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
          ||||||||||||||||||||||||||||||||
Db     601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632


RESULT 4
US-09-488-725A-3192
; Sequence 3192, Application US/09488725A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784FLPCT
;  CURRENT APPLICATION NUMBER: US/09/488,725A
;  CURRENT FILING DATE:  2000-12-22
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7144
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 3192
;   LENGTH: 633
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-488-725A-3192

  Query Match           99.7%;  Score 3233;  DB 18;  Length 632;
  Best Local Similarity  99.7%;  Pred. No. 1.9e-262;
  Matches  630;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60

Qy      61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120

Qy     121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180

Qy     181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240

Qy     241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300

Qy     301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360

Qy     361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420

Qy     421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480

Qy     481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMLQDIKRQPKVVERLLSLPLEYWS 540

Qy     541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db     541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSGSLSQRVRS 600

Qy     601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
          ||||||||||||||||||||||||||||||||
Db     601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632


RESULT 5
US-09-791-537-72082
; Sequence 72082, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 72082
;   LENGTH: 632
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-791-537-72082
```

Query Match 99.6%; Score 3230; DB 21; Length 632;
Best Local Similarity 99.7%; Pred. No. 3.5e-262;
Matches 630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60

Qy         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120

Qy        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSHISFDKTDESLDWDSSLVKTFKLKKR 180

Qy        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240

Qy        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300

Qy        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360

Qy        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420

Qy        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480

Qy        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMLQDIKRQPKVVERLLSLPLEYWS 540

Qy        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600

Qy        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              ||||||||||||||||||||||||||||||||
Db        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
```

RESULT 6
US-09-833-790-413
; Sequence 413, Application US/09833790
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Secrist, Heather
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Indirias, Carol Y.
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.512
; CURRENT APPLICATION NUMBER: US/09/833,790
; CURRENT FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 413
; LENGTH: 632
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-790-413

Query Match 99.6%; Score 3230; DB 22; Length 632;
Best Local Similarity 99.7%; Pred. No. 3.5e-262;
Matches 630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60

Qy         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120

Qy        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSHISFDKTDESLDWDSSLVKTFKLKKR 180

Qy        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240

Qy        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300

Qy        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360

Qy        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420

Qy        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480

Qy        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMLQDIKRQPKVVERLLSLPLEYWS 540

Qy        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600

Qy        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              ||||||||||||||||||||||||||||||||
Db        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
```

RESULT 7
US-09-833-790-427
; Sequence 427, Application US/09833790
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Secrist, Heather
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Indirias, Carol Y.
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.512
; CURRENT APPLICATION NUMBER: US/09/833,790
; CURRENT FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 427
; LENGTH: 570
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-790-427

Query Match 90.0%; Score 2919; DB 22; Length 570;

```
  Best Local Similarity   99.8%;  Pred. No. 4.5e-236;
  Matches  569;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         63 TERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEEQK 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEEQK 60

Qy        123 SALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKREK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKREK 120

Qy        183 RRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVPYW 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVPYW 180

Qy        243 TRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVP 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVP 240

Qy        303 CGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSPMI 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSPMI 300

Qy        363 PSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICSLL 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 PSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICSLL 360

Qy        423 KDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQR 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQR 420

Qy        483 VAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWSQF 542
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        421 VAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMLQDIKRQPKVVERLLSLPLEYWSQF 480

Qy        543 MMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRSTL 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 MMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRSTL 540

Qy        603 TKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              |||||||||||||||||||||||||||||
Db        541 TKNTPRFGSKSKSATNLGRQGNFFASPMLK 570


RESULT 8
US-09-791-537-36304
; Sequence 36304, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36304
;   LENGTH: 628
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-791-537-36304

  Query Match             84.3%;  Score 2734;  DB 21;  Length 628;
  Best Local Similarity   84.4%;  Pred. No. 2e-220;
  Matches  534;  Conservative   47;  Mismatches   46;  Indels    6;  Gaps    3;

Qy          1 MDTMMLNVRNLFEQLVRRVEILSEGNE-VQFIQLAKDFEDFRKKWQRTDHELGKYKDLLM 59
              ||| |:|:  ||||||||:||::|||| ::|||: |||||||||:|||: || |:||||:
Db          1 MDTTMVNLWTLFEQLVRRMEIINEGNESIEFIQVVKDFEDFRKKYQRTNQELEKFKDLLL 60

Qy         60 KAETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSE 119
              |||| |||||||||||||||||||||||||:| |||:||||||::||||||||||||||
Db         61 KAETGRSALDVKLKHARNQVDVEIKRRQRAEAECAKLEQQIQLIRDILMCDTSGSIQLSE 120

Qy        120 EQKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKK 179
              |||||||||||| || :||| |||||||||||||||||||||||||||||||||| ||:||
Db        121 EQKSALAFLNRGQASSGHAGNNRLSTIDESGSILSDISFDKTDESLDWDSSLVKNFKMKK 180

Qy        180 REKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETV 239
              |||||| |||||||||||| |||| ||| ||||||||||||:||||||||||||||||:
Db        181 REKRRSNSRQFIDGPPGPVKKTCSIGSTVDQANESIVAKTTVTVPSDGGPIEAVSTIETL 240

Qy        240 PYWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPES 299
              | ||||| ||:| ||| |||| |||| |||||:||::|||: ||||||||||||||||||
Db        241 PSWTRSRGKSGPLQPVNSDSALNSRPLEPRTDTDNLGTPQNTGGMRLHDFVSKTVIKPES 300

Qy        300 CVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTS 359
              |||||||||||||||||||:|||||||||||||||| ||:|||||||||||||||||| |
Db        301 CVPCGKRIKFGKLSLKCRDCRLVSHPECRDRCPLPCIPPLVGTPVKIGEGMLADFVSQAS 360

Qy        360 PMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAIC 419
              ||||:||| |||||||||||| ||||||||||||||||||:|||||||||||||| ||
Db        361 PMIPAIVVSCVNEIEQRGLTEAGLYRISGCDRTVKELKEKFLKVKTVPLLSKVDDIHVIC 420

Qy        420 SLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIH 479
              |||||||||||||||| ||::|||||||||||||| ||||||| |||||||||||||||
Db        421 SLLKDFLRNLKEPLLTFWLSKAFMEAAEITDEDNSTAAMYQAVSELPQANRDTLAFLMIH 480

Qy        480 LQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYW 539
              ||||:||| |||||:||||||||||||||| |||||||| |||||| ||||||||||||
Db        481 LQRVSQSPDTKMDIANLAKVFGPTIVAHTVPNPDPVTMFQDIKRQLKVVERLLSLPLEYW 540

Qy        540 SQFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVR 599
              :|||||:|||||     :  |  |||:|||:|||||||||||| ||:||| |||||||:
Db        541 NQFMMVDQENIDS----QRGNGNSTPRTPDVKVSLLGPVTTPEFQLVKTPLSSSLSQRLY 596

Qy        600 STLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              : |:|:|||||:|||||||||:|| || :| ||
Db        597 N-LSKSTPRFGNKSKSATNLGQQGKFFPAPYLK 628


RESULT 9
US-09-881-736-4
; Sequence 4, Application US/09881736
; GENERAL INFORMATION:
;  APPLICANT: Glotzer, Michael
;  APPLICANT:  Jantsch-Plunger, Verena
;  APPLICANT:  Romano, Alper
;  APPLICANT:  Mishima, Masanori
;  APPLICANT:  Kaitna, Susanne
;  TITLE OF INVENTION: Cyk-4 polypeptides, DNA molecules encoding them and their use
;  TITLE OF INVENTION:  screening methods
;  FILE REFERENCE: 0652.2260001/EKS/AES
;  CURRENT APPLICATION NUMBER: US/09/881,736
;  CURRENT FILING DATE:  2001-06-18
;  PRIOR APPLICATION NUMBER: EP 00 112 880.0
;  PRIOR FILING DATE: 2000-06-19
;  PRIOR APPLICATION NUMBER: EP 01 110 554.1
;  PRIOR FILING DATE: 2001-04-30
;  PRIOR APPLICATION NUMBER: 60/241,231
;  PRIOR FILING DATE: 2000-10-18
;  PRIOR APPLICATION NUMBER: To be determined
;  PRIOR FILING DATE: 2001-06-13
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 628
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-881-736-4

  Query Match             84.3%;  Score 2734;  DB 22;  Length 628;
  Best Local Similarity   84.4%;  Pred. No. 2e-220;
  Matches  534;  Conservative   47;  Mismatches   46;  Indels    6;  Gaps    3;
```

```
Qy          1 MDTMMLNVRNLFEQLVRRVEILSEGNE-VQFIQLAKDFEDFRKKWQRTDHELGKYKDLLM 59
              ||| |:|:  |||||||||:||::|||| ::|||: |||||||||:|||: || |:|||||:
Db          1 MDTTMVNLWTLFEQLVRRMEIINEGNESIEFIQVVKDFEDFRKKYQRTNQELEKFKDLLL 60

Qy         60 KAETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSE 119
              |||| |||||||||||||||||||||||||:| |||:||||||::|||||||||||||||
Db         61 KAETGRSALDVKLKHARNQVDVEIKRRQRAEAECAKLEQQIQLIRDILMCDTSGSIQLSE 120

Qy        120 EQKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKK 179
              |||||||||||| || :||| ||||||||||||||||||||||||||||||||||||:||
Db        121 EQKSALAFLNRGQASSGHAGNNRLSTIDESGSILSDISFDKTDESLDWDSSLVKNFKMKK 180

Qy        180 REKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETV 239
              |||||| |||||||||||| |||| ||| ||||||||||||:|||||||||||||||:
Db        181 REKRRSNSRQFIDGPPGPVKKTCSIGSTVDQANESIVAKTTVTVPSDGGPIEAVSTIETL 240

Qy        240 PYWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPES 299
              | ||||| |:| ||| |||| |||| ||||||:||::|||||: ||||||||||||||||
Db        241 PSWTRSRGKSGPLQPVNSDSALNSRPLEPRTDTDNLGTPQNTGGMRLHDFVSKTVIKPES 300

Qy        300 CVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTS 359
              |||||||||||||||||||:|||||||||||||||| |:|||||||||||||||||| |
Db        301 CVPCGKRIKFGKLSLKCRDCRLVSHPECRDRCPLPCIPPLVGTPVKIGEGMLADFVSQAS 360

Qy        360 PMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAIC 419
              ||||:||| |||||||||||| ||||||||||||||||||:|||||||||||||||| ||
Db        361 PMIPAIVVSCVNEIEQRGLTEAGLYRISGCDRTVKELKEKFLKVKTVPLLSKVDDIHVIC 420

Qy        420 SLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIH 479
              |||||||||||||||| |::|||||||||||||| ||||||| ||||||||||||||||
Db        421 SLLKDFLRNLKEPLLTFWLSKAFMEAAEITDEDNSTAAMYQAVSELPQANRDTLAFLMIH 480

Qy        480 LQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYW 539
              ||||:||| ||||:||||||||||||| |||||||| |||||| ||||||||||||||||
Db        481 LQRVSQSPDTKMDIANLAKVFGPTIVAHTVPNPDPVTMFQDIKRQLKVVERLLSLPLEYW 540

Qy        540 SQFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVR 599
              :|||||:|||||     :  |  |||:|||:|||||||||||| ||:||| |||||||:
Db        541 NQFMMVDQENIDS----QRGNGNSTPRTPDVKVSLLGPVTTPEFQLVKTPLSSSLSQRLY 596

Qy        600 STLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              : |:|:|||||:|||||||||:|| || :| ||
Db        597 N-LSKSTPRFGNKSKSATNLGQQGKFFPAPYLK 628
```

RESULT 10
US-09-629-469A-14587
; Sequence 14587, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14587
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-629-469A-14587

Query Match 36.4%; Score 1182; DB 20; Length 255;
Best Local Similarity 99.2%; Pred. No. 2.3e-90;
Matches 234; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        397 KEKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIA 456
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 KLKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIA 79

Qy        457 AMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVT 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 AMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVT 139

Qy        517 MSQDIKRQPKVVERLLSLPLEYWSQFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLG 576
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 MLQDIKRQPKVVERLLSLPLEYWSQFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLG 199

Qy        577 PVTTPEHQLLKTPSSSSLSQRVRSTLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 PVTTPEHQLLKTPSSSSLSQRVRSTLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 255
```

RESULT 11
US-60-389-987-1045
; Sequence 1045, Application US/60389987
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465P2
; CURRENT APPLICATION NUMBER: US/60/389,987
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 3025
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1045
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-389-987-1045

Query Match 36.4%; Score 1182; DB 27; Length 255;
Best Local Similarity 99.2%; Pred. No. 2.3e-90;
Matches 234; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        397 KEKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIA 456
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 KLKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIA 79

Qy        457 AMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVT 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 AMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVT 139

Qy        517 MSQDIKRQPKVVERLLSLPLEYWSQFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLG 576
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        140 MLQDIKRQPKVVERLLSLPLEYWSQFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLG 199

Qy        577 PVTTPEHQLLKTPSSSSLSQRVRSTLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 PVTTPEHQLLKTPSSSSLSQRVRSTLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 255


RESULT 12
US-60-412-418-1045
; Sequence 1045, Application US/60412418
; GENERAL INFORMATION:
;  APPLICANT: Ghosh, Soumitra S.
;  APPLICANT:  Fahy, Eoin D.
;  APPLICANT:  Zhang, Bing
;  APPLICANT:  Gibson, Bradford W.
;  APPLICANT:  Taylor, Steven W.
;  APPLICANT:  Glenn, Gary M.
;  APPLICANT:  Warnock, Dale E.
;  TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
;  TITLE OF INVENTION:  IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
;  FILE REFERENCE: 660088.465P3
;  CURRENT APPLICATION NUMBER: US/60/412,418
;  CURRENT FILING DATE:  2002-09-20
;  NUMBER OF SEQ ID NOS: 3025
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1045
;   LENGTH: 255
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-60-412-418-1045

  Query Match             36.4%;  Score 1182;  DB 27;  Length 255;
  Best Local Similarity   99.2%;  Pred. No. 2.3e-90;
  Matches  234;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        397 KEKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIA 456
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 KLKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIA 79

Qy        457 AMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVT 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 AMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVT 139

Qy        517 MSQDIKRQPKVVERLLSLPLEYWSQFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLG 576
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 MLQDIKRQPKVVERLLSLPLEYWSQFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLG 199

Qy        577 PVTTPEHQLLKTPSSSSLSQRVRSTLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 PVTTPEHQLLKTPSSSSLSQRVRSTLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 255


RESULT 13
US-09-758-472-6751
; Sequence 6751, Application US/09758472
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PH001
;  CURRENT APPLICATION NUMBER: US/09/758,472
;  CURRENT FILING DATE:  2001-01-11
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  NUMBER OF SEQ ID NOS: 9632
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6751
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (233)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (238)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (240)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (242)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-758-472-6751

  Query Match             34.0%;  Score 1103.5;  DB 21;  Length 243;
  Best Local Similarity   97.4%;  Pred. No. 8.5e-84;
  Matches  222;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 63

Qy         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 123

Qy        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 183

Qy        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGG 228
              |||||||||||||||||||||||||||||||||| |     |||||
Db        184 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVEKLQ-CCPNDGG 230


RESULT 14
US-10-235-926-6751
; Sequence 6751, Application US/10235926
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PH001C1N
;  CURRENT APPLICATION NUMBER: US/10/235,926
;  CURRENT FILING DATE:  2002-09-06
;  PRIOR APPLICATION NUMBER: 09/758,472
;  PRIOR FILING DATE: 2001-01-11
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  NUMBER OF SEQ ID NOS: 9632
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6751
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (233)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (238)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (240)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (242)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-235-926-6751
```

```
  Query Match            34.0%;  Score 1103.5;  DB 26;  Length 243;
  Best Local Similarity  97.4%;  Pred. No. 8.5e-84;
  Matches  222;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 63

Qy         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 123

Qy        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 183

Qy        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGG 228
              |||||||||||||||||||||||||||||||| |     |||||
Db        184 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVEKLQ-CCPNDGG 230


RESULT 15
PCT-US01-08656-9072
; Sequence 9072, Application PC/TUS0108656
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 21272-066
;  CURRENT APPLICATION NUMBER: PCT/US01/08656
;  CURRENT FILING DATE:  2001-04-16
;  PRIOR APPLICATION NUMBER: 09/522,929
;  PRIOR FILING DATE: 2000-04-18
;  PRIOR APPLICATION NUMBER: 09/770,160
;  PRIOR FILING DATE: 2001-01-26
;  NUMBER OF SEQ ID NOS: 10994
;  SOFTWARE: Custom
; SEQ ID NO 9072
;   LENGTH: 256
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: DOMAIN
;   LOCATION: (7)..(33)
;   OTHER INFORMATION: PROTEIN GTPASE DOMAIN ACTIVATION domain identified by
;   OTHER INFORMATION: eMATRIX, accession number PD00930A, p-value=7.324e-15, raw score
;   OTHER INFORMATION: 25.62
CT-US01-08656-9072

  Query Match            30.0%;  Score 971.5;  DB 1;  Length 256;
  Best Local Similarity  84.0%;  Pred. No: 1.2e-72;
  Matches  205;  Conservative    3;  Mismatches   27;  Indels    9;  Gaps    4;

Qy        357 QTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIH 416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 QTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIH 60

Qy        417 AICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFL 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFL 120

Qy        477 MIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKR-QPKVVERLLSLP 535
              |||||||||||||||||||||||||||||||||||| |   :  ||||||||| |
Db        121 MIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMLQGHQGVQPKVVERLLFLG 180

Qy        536 L-EYWSQFMMVEQE---NIDPLHVIENSNAFSTPQTPDIKV----SLLGPVTTPEHQLLK 587
              | |||    |   |     ||||||||||||||||||||      |    |  | :: |
Db        181 LWEYWEFSFMDGWEQGGTFDPLHVIENSNAFSTPQTPDIKAVPGGGLCVHFTAGEAEIQK 240

Qy        588 TPSS 591
              | |
Db        241 GPPS 244
```

```
Search completed: July  3, 2003, 10:02:52
Job time : 307 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 3, 2003, 09:50:02 ; Search time 88 Seconds
(without alignments)
1479.793 Million cell updates/sec

Title: US-09-881-736-2
Perfect score: 3243
Sequence: 1 MDTMMLNVRNLFEQLVRRVE.....SKSATNLGRQGNFFASPMLK 632

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3243	100.0	632	4	Q9P2W2	Q9p2w2 homo sapien
2	3237	99.8	632	4	Q9H0H5	Q9h0h5 homo sapien
3	3230	99.6	632	4	Q9NWN2	Q9nwn2 homo sapien
4	2919	90.0	570	4	Q9P250	Q9p250 homo sapien
5	2781	85.8	628	4	Q9Bz74	Q9bz74 homo sapien
6	2734	84.3	628	11	Q9WVM1	Q9wvm1 mus musculu
7	1182	36.4	255	4	Q9H9L9	Q9h919 homo sapien
8	864	26.6	625	5	Q9V6U7	Q9v6u7 drosophila
9	862	26.6	625	5	Q9N9Z9	Q9n9z9 drosophila
10	547.5	16.9	681	5	Q9XUS9	Q9xus9 caenorhabdi
11	431	13.3	354	5	Q95VR1	Q95vr1 venturia ca
12	328	10.1	334	11	Q91V57	Q91v57 mus musculu
13	322	9.9	459	4	Q96FB0	Q96fb0 homo sapien
14	320	9.9	299	13	Q92153	Q92153 serinus sp.
15	308	9.5	695	3	O94466	O94466 schizosacch
16	305.5	9.4	1388	5	Q8ST20	Q8st20 caenorhabdi
17	305.5	9.4	1439	5	Q9TY64	Q9ty64 caenorhabdi
18	301	9.3	1261	4	O15463	O15463 homo sapien
19	294	9.1	282	11	Q9D9W2	Q9d9w2 mus musculu
20	291	9.0	970	4	Q9P107	Q9p107 homo sapien
21	290.5	9.0	733	4	Q9BYS7	Q9bys7 homo sapien
22	290.5	9.0	759	4	Q9UJ00	Q9uj00 homo sapien
23	290.5	9.0	814	4	Q9UNA1	Q9una1 homo sapien
24	287.5	8.9	263	11	Q9ER57	Q9er57 mus musculu
25	286.5	8.8	1286	4	Q9P2C3	Q9p2c3 homo sapien
26	285.5	8.8	1165	4	Q92619	Q92619 homo sapien
27	281.5	8.7	889	4	Q12844	Q12844 homo sapien
28	281.5	8.7	936	4	Q12843	Q12843 homo sapien
29	275.5	8.5	210	11	Q9D9B3	Q9d9b3 mus musculu
30	270.5	8.3	568	13	Q98935	Q98935 gallus gall
31	270.5	8.3	1397	4	Q9NTG2	Q9ntg2 homo sapien
32	270.5	8.3	2548	4	Q9UNJ2	Q9unj2 homo sapien
33	269.5	8.3	735	4	Q9P227	Q9p227 homo sapien
34	267.5	8.2	2626	11	Q9Z1N3	Q9z1n3 rattus norv
35	266.5	8.2	523	11	Q9CU46	Q9cu46 mus musculu
36	265	8.2	494	5	Q9W4A9	Q9w4a9 drosophila
37	260.5	8.0	802	4	O60890	O60890 homo sapien
38	258.5	8.0	714	3	O13384	O13384 cochliobolu
39	258.5	8.0	786	11	Q99MT3	Q99mt3 mus musculu
40	258	8.0	903	4	O14560	O14560 homo sapien
41	256.5	7.9	632	4	Q96EV3	Q96ev3 homo sapien
42	256.5	7.9	714	3	O13390	O13390 cochliobolu
43	256.5	7.9	786	4	Q96S75	Q96s75 homo sapien
44	255.5	7.9	245	11	Q99LW5	Q99lw5 mus musculu
45	254	7.8	683	11	Q99MT4	Q99mt4 mus musculu

ALIGNMENTS

RESULT 1
Q9P2W2
ID Q9P2W2 PRELIMINARY; PRT; 632 AA.
AC Q9P2W2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE GTPase activating protein.
GN MGCRACGAP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20435340; PubMed=10979956;
RA Kawashima T., Hirose K., Satoh T., Kaneko A., Ikeda Y., Kaziro Y.,
RA Nosaka T., Kitamura T.;
RT "MgcRacGAP is involved in the control of growth and differentiation of
RT hematopoietic cells.";
RL Blood 96:2116-2124(2000).
DR EMBL; AB030251; BAA90247.1; -.
DR HSSP; Q07960; 1RGP.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR000198; RhoGAP.
DR Pfam; PF00130; DAG_PE-bind; 1.
DR Pfam; PF00620; RhoGAP; 1.
DR SMART; SM00109; C1; 1.
DR SMART; SM00324; RhoGAP; 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.
DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
SQ SEQUENCE 632 AA; 71000 MW; BC2B75E5A8739E2B CRC64;

Query Match 100.0%; Score 3243; DB 4; Length 632;
Best Local Similarity 100.0%; Pred. No. 1.6e-213;
Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60

```
Db          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60

Qy         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120

Qy        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180

Qy        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240

Qy        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300

Qy        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360

Qy        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420

Qy        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480

Qy        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540

Qy        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600

Qy        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              ||||||||||||||||||||||||||||||||
Db        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
```

```
RESULT 2
Q9H0H5
ID   Q9H0H5      PRELIMINARY;      PRT;   632 AA.
AC   Q9H0H5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical 71.0 kDa protein.
GN   DKFZP434C011.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RX   MEDLINE=21154917; PubMed=11230166;
RA   Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA   Ansorge W., Boecher M., Bloecker H., Bauersachs S., Blum H.,
RA   Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA   Mewes H.W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA   Wambutt R., Korn B., Klein M., Poustka A.;
RT   "Towards a Catalog of Human Genes and Proteins: Sequencing and
RT   Analysis of 500 Novel Complete Protein Coding Human cDNAs.";
RL   Genome Res. 11:422-435(2001).
DR   EMBL; AL136794; CAB66728.1; -.
DR   HSSP; Q07960; 1RGP.
DR   InterPro; IPR002219; DAG_PE-bind.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00109; C1; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   632 AA;  71026 MW;  032B7DF9CEA8F39D CRC64;

  Query Match             99.8%;  Score 3237;  DB 4;  Length 632;
  Best Local Similarity   99.8%;  Pred. No. 4e-213;
  Matches  631;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60

Qy         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120

Qy        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180

Qy        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240

Qy        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300

Qy        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360

Qy        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420

Qy        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480

Qy        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMLQDIKRQPKVVERLLSLPLEYWS 540

Qy        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600

Qy        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              ||||||||||||||||||||||||||||||||
Db        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
```

```
RESULT 3
Q9NWN2
ID   Q9NWN2      PRELIMINARY;      PRT;   632 AA.
AC   Q9NWN2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   CDNA FLJ20726 fis, clone HEP13735.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA   Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Obayashi M.,
```

```
RA   Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK000733; BAA91347.1; -.
DR   HSSP; Q07960; 1RGP.
DR   InterPro; IPR002219; DAG_PE-bind.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00109; C1; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
SQ   SEQUENCE   632 AA;  71048 MW;  7DC06CB725BBAAA4 CRC64;

  Query Match           99.6%;  Score 3230;  DB 4;  Length 632;
  Best Local Similarity  99.7%;  Pred. No. 1.2e-212;
  Matches  630;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60

Qy         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120

Qy        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSHISFDKTDESLDWDSSLVKTFKLKKR 180

Qy        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240

Qy        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300

Qy        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360

Qy        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420

Qy        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480

Qy        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMLQDIKRQPKVVERLLSLPLEYWS 540

Qy        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600

Qy        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              ||||||||||||||||||||||||||||||||
Db        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632


RESULT 4
Q9P250
ID   Q9P250      PRELIMINARY;      PRT;   570 AA.
AC   Q9P250;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   KIAA1478 protein (Fragment).
GN   KIAA1478.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20277482; PubMed=10819331;
RA   Nagase T., Kikuno R., Ishikawa K., Hirosawa M., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human
RT   genes.XVII.The complete sequences of 100 new cDNA clones from brain
RT   which code for large proteins in vitro.";
RL   DNA Res. 7:143-150(2000).
DR   EMBL; AB040911; BAA96002.1; -.
DR   HSSP; Q07960; 1RGP.
DR   InterPro; IPR002219; DAG_PE-bind.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00109; C1; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   570 AA;  63468 MW;  A03F36C0AAC65485 CRC64;

  Query Match           90.0%;  Score 2919;  DB 4;  Length 570;
  Best Local Similarity  99.8%;  Pred. No. 1.9e-191;
  Matches  569;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         63 TERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEEQK 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEEQK 60

Qy        123 SALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKREK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKREK 120

Qy        183 RRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVPYW 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVPYW 180

Qy        243 TRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVP 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVP 240

Qy        303 CGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSPMI 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSPMI 300

Qy        363 PSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICSLL 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 PSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICSLL 360

Qy        423 KDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQR 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQR 420

Qy        483 VAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWSQF 542
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        421 VAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMLQDIKRQPKVVERLLSLPLEYWSQF 480

Qy        543 MMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRSTL 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 MMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRSTL 540

Qy        603 TKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              ||||||||||||||||||||||||||||||
Db        541 TKNTPRFGSKSKSATNLGRQGNFFASPMLK 570


RESULT 5
Q9BZ74
ID   Q9BZ74      PRELIMINARY;      PRT;   628 AA.
```

```
AC   Q9BZ74;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   FKSG42.
GN   FKSG42.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wang Y.-G., Gong L.;
RT   "Characterization of FKSG42, a novel gene located on human chromosome
RT   12.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF334184; AAG50293.1; -.
DR   HSSP; Q07960; 1RGP.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00324; RhoGAP; 1.
SQ   SEQUENCE   628 AA;  70013 MW;  A142855A613154CC CRC64;

  Query Match           85.8%;  Score 2781;  DB 4;  Length 628;
  Best Local Similarity   87.0%;  Pred. No. 6.1e-182;
  Matches  550;  Conservative   31;  Mismatches   47;  Indels    4;  Gaps    2;

Qy          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60
              ||| |||:||||||||  ||| ||||:||||| |||||| |||:|||:||||||||:|
Db          1 MDTTMLNMRNLFEQLVRGAEILKEGNELQFIQLVKDFEDFFKKWERTDYELGKYKDLLIK 60

Qy         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120
              |||||||||||||| |||| :|:||:||||:|||||:||||| || :|| ||||||| |
Db         61 AETERSALDVKLKHAHNQVDADIQRRRRAEANCEKLEQQIQLIGEMFLCDISGSIQLSGE 120

Qy        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180
              |||||||||||||||||||||  |||||||| |||||||||| ||||||||| |:|
Db        121 QKSALAFLNRGQPSSSNAGNKRLSISDESGSILSGISFDKTDESLDCDSSLVKTFKRKQR 180

Qy        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240
              ||| |:|||||||| ||||:|||||| |||||||||||||||| |||||:|| ||||||
Db        181 EKRCSSSRQFVDGPSGPVKETRSIGSTVDQGNESIVAKTTVTVPKDGGPIKAVPTIETVP 240

Qy        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300
              |||||:|||||||||||| ||||||||||||||:||||||| || |:||||||||||||
Db        241 YWTRSQRKTGTLQPWNSDPTLNSRQLEPRTETDNVGTPQSNRGMCLNDFVSKTVIKPEFY 300

Qy        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360
              |||||:||||||||||:|| |:| ||||| |||||||||||||||||||||||||||| |
Db        301 VPCGKQIKFGKLSLKCQDCCVISRPECRDHCPLPCIPTLIGTPVKIGEGMLADFVSQTFP 360

Qy        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420
              ||| |||||||||||||||||||| ||| ||| ||||||||||||:||||||||||||
Db        361 MIPRIVVHCVNEIEQRGLTETGLYRILGCDSTVK-LKEKFLRVKTVPLVSKVDDIHAICS 419

Qy        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480
              |||||||| |||||||||:|||||||||||||||:||||||||||||||||| ||||||
Db        420 LLKDFLRNFKEPLLTFRLNKAFMEAAEITDEDNSISAMYQAVGELPQANRDTLVFLMIHL 479

Qy        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540
              |||||||:|||:|||||:||| |||||||||:|||| |||| |||||||| ||||||||
Db        480 QRVAQSPYTKMNVANLAEVFGSTIVAHAVPNPEPVTMLQDIKCQPKVVERLPSLPLEYWS 539

Qy        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600
              || :|||:||||||| |||||||   |||||||||||||||:||||||||||||| | |
Db        540 QFSIVEQKNIDPLHVTENSNAFS---TPDIKVSLLGPVTTPDHQLLKTPSSSSLSQGVCS 596

Qy        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              |||||||||||| |||||||||| ||| |||
Db        597 TLTKNTPRFGSKGKSATNLGRQGNLFASSMLK 628


RESULT 6
```

```
Q9WVM1
ID   Q9WVM1      PRELIMINARY;      PRT;   628 AA.
AC   Q9WVM1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   RAC GTPase-activating protein (GTPase activating protein ID-GAP)
DE   (MGCRACGAP variant beta) (MGCRACGAP variant alpha) (RAC
DE   GTPase-activating protein 1).
GN   RACGAP1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99249188; PubMed=10235109;
RA   Wooltorton E.J., Haliotis T., Mueller C.R.;
RT   "Identification and characterization of a transcript for a novel Rac
RT   GTPase-activating protein in terminally differentiating 3T3-L1
RT   adipocytes.";
RL   DNA Cell Biol. 18:265-273(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Kawashima T., Kitamura T.;
RT   "Mus musculus cDNA for GTPase activating protein MgcRacGAP, complete
RT   cds.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CD-1;
RA   Van de Putte T., Zwijsen A., Lonnoy O., Rybin V., Cozijnsen M.,
RA   Francis A., Baekelandt V., Kozak C.A., Zerial M., Huylebroeck D.;
RT   "Mice with a homozygous gene trap vector insertion in MgcRacGAP die
RT   during pre-implantation development.";
RL   Mech. Dev. 0:0-0(2001).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF079974; AAD40487.1; -.
DR   EMBL; AB030252; BAA90248.1; -.
DR   EMBL; AF212321; AAG43540.1; -.
DR   EMBL; AF212320; AAG43539.1; -.
DR   EMBL; BC010715; AAH10715.1; -.
DR   HSSP; Q07960; 1RGP.
DR   MGD; MGI:1349423; Racgap1.
DR   InterPro; IPR002219; DAG_PE-bind.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00109; C1; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
SQ   SEQUENCE   628 AA;  70158 MW;  8D1B9DEC3CE057BE CRC64;

  Query Match           84.3%;  Score 2734;  DB 11;  Length 628;
  Best Local Similarity   84.4%;  Pred. No. 1e-178;
  Matches  534;  Conservative   47;  Mismatches   46;  Indels    6;  Gaps    3;

Qy          1 MDTMMLNVRNLFEQLVRRVEILSEGNE-VQFIQLAKDFEDFRKKWQRTDHELGKYKDLLM 59
              ||| |:|:  |||||||:||::|||| ::|||: |||||||||:|||: || |:||||:
Db          1 MDTTMVNLWTLFEQLVRRMEIINEGNESIEFIQVVKDFEDFRKKYQRTNQELEKFKDLLL 60

Qy         60 KAETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSE 119
              |||| |||||||||||||||||||||||:| |||:||||||::||||||||||||
Db         61 KAETGRSALDVKLKHARNQVDVEIKRRQRAEAECAKLEQQIQLIRDILMCDTSGSIQLSE 120

Qy        120 EQKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKK 179
              |||||||||||| || :||| |||||||||||||||||||||||||||||| ||:||
Db        121 EQKSALAFLNRGQASSGHAGNNRLSTIDESGSILSDISFDKTDESLDWDSSLVKNFKMKK 180
```

```
Qy        180 REKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETV 239
              ||||||  ||||:||||||||||  ||||  |||  ||||||||||||||:||||||||||||:
Db        181 REKRRSNSRQFIDGPPGPVKKTCSIGSTVDQANESIVAKTTVTVPSDGGPIEAVSTIETL 240

Qy        240 PYWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPES 299
              | |||||  |:|  |||  ||||  ||||  |||||:||::||||: ||||||||||||||||||
Db        241 PSWTRSRGKSGPLQPVNSDSALNSRPLEPRTDTDNLGTPQNTGGMRLHDFVSKTVIKPES 300

Qy        300 CVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTS 359
              ||||||||||||||||||||:|||||||||||||||  |:|||||||||||||||||  |
Db        301 CVPCGKRIKFGKLSLKCRDCRLVSHPECRDRCPLPCIPPLVGTPVKIGEGMLADFVSQAS 360

Qy        360 PMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAIC 419
              ||||:|||  ||||||||||| ||||||||||||||||||:|||||||||||||  ||
Db        361 PMIPAIVVSCVNEIEQRGLTEAGLYRISGCDRTVKELKEKFLKVKTVPLLSKVDDIHVIC 420

Qy        420 SLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIH 479
              ||||||||||||||||  |::|||||||||||||||  |||||||  ||||||||||||||||
Db        421 SLLKDFLRNLKEPLLTFWLSKAFMEAAEITDEDNSTAAMYQAVSELPQANRDTLAFLMIH 480

Qy        480 LQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYW 539
              ||||:|||  ||||:|||||||||||||||  |||||||||  ||||||  ||||||||||||||
Db        481 LQRVSQSPDTKMDIANLAKVFGPTIVAHTVPNPDPVTMFQDIKRQLKVVERLLSLPLEYW 540

Qy        540 SQFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVR 599
              :|||||:|||||     :  |  |||:|||:||||||||||||  ||:|||  |||||||:
Db        541 NQFMMVDQENIDS----QRGNGNSTPRTPDVKVSLLGPVTTPEFQLVKTPLSSSLSQRLY 596

Qy        600 STLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              : |:|:|||||:|||||||||:|| || :| ||
Db        597 N-LSKSTPRFGNKSKSATNLGQQGKFFPAPYLK 628


RESULT 7
Q9H9L9
ID   Q9H9L9      PRELIMINARY;      PRT;   255 AA.
AC   Q9H9L9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   CDNA FLJ12664 fis, clone NT2RM4002226, weakly similar to GTPase
DE   activating protein rotund.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA   Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA   Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA   Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA   Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA   Ninomiya K., Iwayanagi T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK022726; BAB14206.1; -.
DR   HSSP; Q07960; 1RGP.
DR   InterPro; IPR001092; HLH_basic.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
SQ   SEQUENCE   255 AA;  28488 MW;  B59D7B59385821BD CRC64;

  Query Match           36.4%;  Score 1182;  DB 4;  Length 255;
  Best Local Similarity   99.2%;  Pred. No. 3.9e-73;
  Matches  234;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        397 KEKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIA 456
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 KLKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIA 79

Qy        457 AMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVT 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 AMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVT 139

Qy        517 MSQDIKRQPKVVERLLSLPLEYWSQFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLG 576
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 MLQDIKRQPKVVERLLSLPLEYWSQFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLG 199

Qy        577 PVTTPEHQLLKTPSSSSLSQRVRSTLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 PVTTPEHQLLKTPSSSSLSQRVRSTLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 255


RESULT 8
Q9V6U7
ID   Q9V6U7      PRELIMINARY;      PRT;   625 AA.
AC   Q9V6U7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   CG13345 protein.
GN   RACGAP50C OR CG13345.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003816; AAF58324.1; -.
DR   FlyBase; FBgn0033881; RacGAP.
DR   InterPro; IPR002219; DAG_PE-bind.
```

```
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00109; C1; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
SQ   SEQUENCE   625 AA;  69762 MW;  535F2CF986AA8F03 CRC64;

  Query Match             26.6%;  Score 864;  DB 5;  Length 625;
  Best Local Similarity   35.1%;  Pred. No. 7.5e-51;
  Matches  228;  Conservative  110;  Mismatches  225;  Indels   86;  Gaps   21;

Qy          4 MMLNVRNLFEQLVRRVEILSEGN-EVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMKAE 62
              | |:    |: | | :::|::|  | :|::  : || : :|      |  : :: | |:
Db          1 MALSALASFDDLRRCMQVLTDGTPEEEFLRFLRMFEQYHEKCAGYAAETARIQNELDKSL 60

Qy         63 TERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEEQK 122
              |:   |: || |||  :|:||| |::|| : : :| :|  : ::|  : :    |: | :
Db         61 TKMGDLEGKLFHARRIIDMEIKARRQAEHERDAMESKIMAVADLLRHERN----LNNETR 116

Qy        123 SALAFLNRGQPSSS-----NA--GNKRLSTIDESGSILSDISFDKT-DESLDWDSSLVKT 174
              ||||:   |||      ||   :|    |: :|||:|||:|   : |: ||     |:|
Db        117 DKLAFLHT-LPSSRKRKSLNAVREDKSYGDINSTGSLLSDLSITHSEDDFLD-----VRT 170

Qy        175 FKLKKREKRRSTSRQFV-----------------DGPPGPVKKTR--SIGSAVD----- 209
              |   || | |  :  :                      |     |:|  |:|  |:
Db        171 SK-SWREHRPSLPKNQIPSVGNKRSRLSTGLNGSMSGTTPTTGKSRRSSVGIGVEQHTVD 229

Qy        210 --QGNESIVAKTTVTVPNDG-GPIEAVSTIETVPYWTRSRRKTGTLQPWNSDSTLNSRQL 266
              || |  | | ||:| || | | | ||||::|   : |    |       | |
Db        230 VGQGAERFCATTKVTIPQDGQGVIRAESTIESLPVIAGNERIGDGLSSTPRRSVLKEATA 289

Qy        267 EPRTETDSV--------GTP-QSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLSLKCR 317
              | |  :::        ||| |    || | |  || :: ::|| | |||:|| : |:||
Db        290 PPLTPVNAMAPHVVAESGTPLQHRPLMRNHTFSQKTFLRGDNCVQCQKRIRFGAVGLRCR 349

Qy        318 DCRVVSHPECRDRCPLPCIPTLIGTP-VKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQR 376
              || |  | :||    : :|:|   |||  |   | : ||    :||||:::|||||||| |
Db        350 DCPVRCHIDCRYLLTVSCVPQ-TGTPTTKTMTGYVTDFAPSIAPMIPALIVHCVNEIEAR 408

Qy        377 GLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTF 436
              |||| ||||:|  :|  | |||:||| |  | |    ||: :|  :|||||:| |||:
Db        409 GLTEVGLYRLSSSEREYKALKEQFLRGKATPHLGNT-DIYVLCCCVKDFLRSLTEPLIPT 467

Qy        437 RLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTKMDVANL 496
              : |  | :  |   :  : ::| :|||||||||||||:| ||:|| |   | : |:
Db        468 SQWKDFANAVQNPDTKTAQDMLVKSVKQLPQANRDTLAFLILHFQRIAQCPVVLMPIDNI 527

Qy        497 AKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWSQFMMVEQENIDPLHVI 556
              : :|||||| :: |:||   :  ::  | :|:: || ||: :| |:::     |||
Db        528 SLIFGPTIVGYSTPDPDQHAIYTEVFTQKQVMKALLELPVSFWEQYIV-----IDPTR-- 580

Qy        557 ENSNAFSTPQT-----PDIKVSLLGPVTTP-------EHQLLKTPSSSS 593
                    || |     |  |  ||    ||       : :   || :|:
Db        581 -------TPATVIKRVPSNKNDLLSLYATPFKGGTIKKRKFYGTPPASA 622


RESULT 9
Q9N9Z9
ID   Q9N9Z9      PRELIMINARY;      PRT;   625 AA.
AC   Q9N9Z9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   AcGAP protein.
GN   RACGAP50C OR CG13345.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sotillos S., Campuzano S.;
RT   "Cooperation between Ras and Rac signalling pathways in Drosophila
RT   wing development.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ251502; CAB96203.1; -.
DR   FlyBase; FBgn0033881; RacGAP.
DR   InterPro; IPR002219; DAG_PE-bind.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00109; C1; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
SQ   SEQUENCE   625 AA;  69763 MW;  E3892199534D19D0 CRC64;

  Query Match             26.6%;  Score 862;  DB 5;  Length 625;
  Best Local Similarity   35.2%;  Pred. No. 1e-50;
  Matches  229;  Conservative  109;  Mismatches  223;  Indels   90;  Gaps   22;

Qy          4 MMLNVRNLFEQLVRRVEILSEGN-EVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMKAE 62
              | |:    |: | | :::|::|  | :|::  : || : :|      |  : :: | |:
Db          1 MALSALASFDDLRRCMQVLTDGTPEEEFLRFLRMFEQYHEKCAGYAAETARIQNELDKSL 60

Qy         63 TERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEEQK 122
              |:   |: || |||  :|:||| |::|| : : :| :|  : ::|  : :    |: | :
Db         61 TKMGDLEGKLFHARRIIDMEIKARRQAEHERDAMESKIMAVADLLRHERN----LNNETR 116

Qy        123 SALAFLNRGQPSSSNAGNKRLSTI--DES-------GSILSDISFDKT-DESLDWDSSLV 172
              ||||:   |||      | |: :  |||        ||:|||:|   : |: ||     |
Db        117 DKLAFLHT-LPSSRK--RKSLNAVREDESYGDINSTGSLLSDLSITHSEDDFLD-----V 168

Qy        173 KTFKLKKREKRRSTSRQFV-----------------DGPPGPVKKTR--SIGSAVD--- 209
              :| |   || | |  :  :                      |     |:|  |:|  |:
Db        169 RTSK-SWREHRPSLPKNQIPSVGNKRSRLSTGLNGSMSGTTPTTGKSRRSSVGIGVEQHT 227

Qy        210 ----QGNESIVAKTTVTVPNDG-GPIEAVSTIETVPYWTRSRRKTGTLQPWNSDSTLNSR 264
                  || |   | | ||:| || | | | ||||::|    : |    |       | |
Db        228 VDVGQGAERFCATTKVTIPQDGQGVIRAESTIESLPVIAGNERIGDGLSSTPRRSVLKEA 287

Qy        265 QLEPRTETDSV--------GTP-QSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLSLK 315
                | |  :::        ||| |    || | |  || :: ::|| | |||:|| : |:
Db        288 TAPPLTPVNAMAPHVVAESGTPLQHRPLMRNHTFSQKTFLRGDNCVQCQKRIRFGAVGLR 347

Qy        316 CRDCRVVSHPECRDRCPLPCIPTLIGTP-VKIGEGMLADFVSQTSPMIPSIVVHCVNEIE 374
              |||| |  | :||    : |:|   |||  |   | : ||    :||||:::|||||||
Db        348 CRDCPVRCHIDCRYLLTVSCVPQ-TGTPTTKTMTGYVTDFAPSIAPMIPALIVHCVNEIE 406

Qy        375 QRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLL 434
               ||||| ||||:|  :|  | |||:||| |  | |    ||: :|  :|||||:| |||:
Db        407 ARGLTEVGLYRLSSSEREYKALKEQFLRGKATPHLGNT-DIYVLCCCVKDFLRSLTEPLI 465

Qy        435 TFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTKMDVA 494
                  : |  | :  |   :  : ::| :|||||||||||||:| ||:|| |   | : |:
Db        466 PTSQWKDFANAVQNPDTKTAQDMLVKSVKQLPQANRDTLAFLILHFQRIAQCPVVLMPID 525

Qy        495 NLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWSQFMMVEQENIDPLH 554
              |:: :|||||| :: |:||   :  ::  | :|:: || ||: :| |:::     |||
Db        526 NISLIFGPTIVGYSTPDPDQHAIYTEVFTQKQVMKALLELPVSFWEQYIV-----IDPTR 580

Qy        555 VIENSNAFSTPQT-----PDIKVSLLGPVTTP-------EHQLLKTPSSSS 593
                    || |     |  |  ||    ||       : :   || :|:
Db        581 ---------TPATVIKRVPSNKNDLLSLYATPFKGGTIKKRKFYGTPPASA 622


RESULT 10
Q9XUS9
ID   Q9XUS9      PRELIMINARY;      PRT;   681 AA.
AC   Q9XUS9;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
```

```
DE   K08E3.6 protein.
GN   K08E3.6.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   McMurray A.A.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z81568; CAB04593.1; -.
DR   InterPro; IPR002219; DAG_PE-bind.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00109; C1; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
SQ   SEQUENCE   681 AA;  76314 MW;  B3DD74F4CC51B685 CRC64;

  Query Match             16.9%;  Score 547.5;  DB 5;  Length 681;
  Best Local Similarity   26.0%;  Pred. No. 3.7e-29;
  Matches  181;  Conservative  109;  Mismatches  284;  Indels  121;  Gaps   22;

Qy          7 NVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKW-------QRTDHELGKYKDLLM 59
              | |::|  ::       :  ::    |  : |  || |       :| : :: : :: |
Db         14 NSRHIFNMILNSQRPQFDIKDIGMFHLIDEIERLRKLWKDSEESKKRLNADMREAEEALA 73

Qy         60 KAETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSE 119
              ||  : :  |: :|  :  :   ::  :  : |    | : : ::: :      |  |::
Db         74 KARKKLAMFDIDVKDTQKHLRALMEENKALKLDLNVYETREKQLKDAMKNGIFNS--LTK 131

Qy        120 EQKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWD---SSLVKTFK 176
              | :    ||:          :||:       :: |   |: |  :|:|    |  :
Db        132 EDRDQFKFLHE---PLVRTYSKRVQ--QRHPHLMEDTQDDEDDSEVDYDETGDSFEEVIH 186

Qy        177 LKK-REKRRSTS-----------------------------RQFVDGPPG----PVKK 200
              |:  || |||::                           :|  |      | |:
Db        187 LRNGREVRRSSAAGNAVGGKRRSASAHAITAAANSKRSRSRVMTATIDEEPNEGGTPPKR 246

Qy        201 TRSIGSAVDQGNESIVAKTTVT------VPNDGGP--------------IEAVSTIETVP 240
              |  ||   |   :    || |        |   |                :: : :  |
Db        247 CRDDGSTPHQEMTTTTTTTTTTTIHNSRAQNQDPPRVSLHRQLTRRSLSCGSIPSCDQTP 306

Qy        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300
                |  :    |      : |||: | |:  |   : || : :  || | |:   :     |
Db        307 GQTTNNIGLGMSSAILTKSTLDIRTLKRGTPAWTNGTTR-DIAMRPHTFIEAGIKAMRKC 365

Qy        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIP----------TLIGTPVKIGEGM 350
              |   :|    |:|||||  | |  | ::  |||||          | |      ||
Db        366 DKCATALKLA-TSMKCRDCHQVVHRSCCNKLHLPCIPRPKTMMTPKSALRGAKPGAGEFR 424

Qy        351 LADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLS 410
              |·|| :   ||||: |:|||  :| ||||: |:||: |  |||  | :: || |||| :
Db        425 LQDFCTSAKPMIPAAVIHCVVALEARGLTQEGIYRVPGQVRTVNVLLDE-LRSKTVPNVG 483

Qy        411 KVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEI--TDEDNSIAAMYQAVGELPQA 468
               : |:  |   || |||:||:||:     :  : || :  || ||   |: : : |||||
Db        484 -LHDVEVITDTLKRFLRDLKDPLIPRTSRQELIVAANLYSTDPDNGRLALNRVICELPQA 542

Qy        469 NRDTLAFLMIHLQRV-AQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQ-------D 520
              ||||||:| || ::| |||   ||:   :|::  | :: |       ||  ||       |
Db        543 NRDTLAYLFIHWRKVIAQSSRNKMNCEAMARMVAPAVMGH------PVKQSQSQAIAGRD 596

Qy        521 IKRQPKVVERLLSLPLEYWSQFMMVEQENIDPLHVIENSNAFSTPQTPD----IKVSLLG 576
                : :  |      || :|:           |   ||   | :  |       |:||
Db        597 ATDCHRAMTALFEFDDVYWQRFLGTSA-------VSMASNQIETARHQDNFALCDRSILG 649

Qy        577 PVTTPEHQLLKTPSSSSLSQRVRSTLTKNTPRFGS 611
              ||||       :|::  |::   :|  :     ||
Db        650 PVTT-------SPATPLLARSANATRARGAHLLGS 677


RESULT 11
Q95VR1
ID   Q95VR1      PRELIMINARY;      PRT;   354 AA.
AC   Q95VR1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Virus-like particle protein VLP2.
OS   Venturia canescens.
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
OC   Ichneumonoidea; Ichneumonidae; Campopleginae; Venturia.
OX   NCBI_TaxID=32260;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Reineke A., Asgari S., Ma G., Beck M., Schmidt O.;
RT   "Sequence analysis and expression of a virus-like particle protein,
RT   VLP2, from the parasitic wasp Venturia canescens.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF410773; AAL02015.1; -.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00620; RhoGAP; 1.
SQ   SEQUENCE   354 AA;  40074 MW;  7DC0949724316449 CRC64;

  Query Match             13.3%;  Score 431;  DB 5;  Length 354;
  Best Local Similarity   37.4%;  Pred. No. 1.4e-21;
  Matches  105;  Conservative   56;  Mismatches  106;  Indels   14;  Gaps    6;

Qy        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420
              ||| :::||:| :| ||:|| ||||:||     : | ::||:      ||:  ||  | |
Db          1 MIPPLIIHCINAVELRGMTELGLYRLSGNKDRYEFLIQQFLKNDGALDLSEY-DIPTITS 59

Qy        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480
               || |||:| |||:|: |   |:||| : |  :    :   | :|||| ||||:||:||
Db         60 ALKQFLRSLSEPLVTYALRDTFIEAAMLPDARDRDVELGSVVMQLPQANLDTLAYLMLHL 119

Qy        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540
              |||::|   ||  :|||||||||:| |:  |  | ::  : | |  ||| ||::| |:|:
Db        120 QRVSESEECKMPKSNLAKVFGPTLVGHSTGNLTPQSLIDEPKYQAAVVEALLNMPEEFWA 179

Qy        541 QFM---MVEQENIDPLHVIENSNAFS-TPQTPDIK----VSLLGPVTTPEHQLLKTPSSS 592
              :|:    ::|  :  |   ::|   |   : |:::       | |:: |      |  |||
Db        180 KFIDPRKLQQRRLSELTRSQSSTQLSQLSRRPNLQGLSGSHALTPLSRPTEAAFKKRSSS 239

Qy        593 ---SLSQRVRSTLTKNTPRFGSKS--KSATNLGRQGNFFAS 628
                  ||:     : :: | : : :      | |:    :||
Db        240 DPEDLSRYGNEEIDQDDPVYENVNLRSQTVNQGQPELLYAS 280


RESULT 12
Q91V57
ID   Q91V57      PRELIMINARY;      PRT;   334 AA.
AC   Q91V57;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   N-chimaerin (Unknown) (Protein for MGC:19150) (RIKEN cDNA 1700112L09
DE   gene).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ILS, AND ISS;
```

```
RA   Ehringer M.A., Thompson J., Conroy O., Xu Y., Yang F., Canniff J.,
RA   Beeson M., Gordon L., Bennett B., Johnson T.E., Sikela J.M.;
RT   "High-Throughput Sequence Identification of Gene Coding Variants
RT   within Alcohol-Related QTLs.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=COLON;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=EYE;
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF332070; AAK56098.1; -.
DR   EMBL; BC010825; AAH10825.1; -.
DR   EMBL; AF332069; AAK56097.1; -.
DR   EMBL; BC024796; AAH24796.1; -.
DR   EMBL; BC025023; AAH25023.1; -.
DR   InterPro; IPR002219; DAG_PE-bind.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
SQ   SEQUENCE   334 AA;  38121 MW;  9EC5DA66D58A36C4 CRC64;

  Query Match           10.1%;  Score 328;  DB 11;  Length 334;
  Best Local Similarity   32.4%;  Pred. No. 1.4e-14;
  Matches   91;  Conservative   46;  Mismatches  118;  Indels   26;  Gaps    8;

Qy        259 STLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLS--LKC 316
              :||   :  |: |             ::|:|  |   |  | |  |     :| ::  :||
Db         66 ATLKENEQIPKYE------------KVHNFKVHTFRGPHWCEYCA-NFMWGLIAQGVKC 111

Qy        317 RDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQR 376
               || :  | :|    |  | | |     |:   |  |       | :|  |: ||| |
Db        112 ADCGLNVHKQCSKMVPNDCKPDLKHVK-KVYSCDLTTLVKAHITKRPMVVDMCIREIESR 170

Qy        377 GLTETGLYRISGCDRTVKELKEKFLR--VKTVPLLSKVDDIHAICSLLKDFLRNLKEPLL 434
              ||   ||||:||    ::::|  | |   |    ::  :||: |   || : |:|  ||:
Db        171 GLNSEGLYRVSGFSDLIEDVKMAFDRDGEKADISVNMYEDINIITGALKLYFRDLPIPLI 230

Qy        435 TFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTK---M 491
              |:    |:|:|:| | |  :  :::|:  || |: :|| :|| ||:||  : | |   |
Db        231 TYDAYPKFIESAKIMDPDEQLETLHEALRSLPPAHCETLRYLMAHLKRV--TLHEKENLM 288

Qy        492 DVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLL 532
              ||  |||||::   |  ||:   ||: |  ||| |:
Db        289 SAENLGIVFGPTLMRS--PELDPMAALNDIRYQRLVVELLI 327


RESULT 13
Q96FB0
ID   Q96FB0      PRELIMINARY;      PRT;   459 AA.
AC   Q96FB0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Similar to chimerin (chimaerin) 1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=EYE;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC011393; AAH11393.1; -.
DR   InterPro; IPR002219; DAG_PE-bind.
DR   InterPro; IPR000198; RhoGAP.
DR   InterPro; IPR000980; SH2.
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   Pfam; PF00017; SH2; 1.
DR   ProDom; PD000093; SH2; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   459 AA;  53172 MW;  04C4CC9BCC611389 CRC64;

  Query Match           9.9%;  Score 322;  DB 4;  Length 459;
  Best Local Similarity   32.0%;  Pred. No. 5.7e-14;
  Matches   90;  Conservative   48;  Mismatches  117;  Indels   26;  Gaps    8;

Qy        259 STLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLS--LKC 316
              :||   :  |: |             ::|:|  |   |  | |  |     :| ::  :||
Db        191 ATLKENEQIPKYE------------KIHNFKVHTFRGPHWCEYCA-NFMWGLIAQGVKC 236

Qy        317 RDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQR 376
               || :  | :|    |  | | |     |:   |  |   :  | :|  |: ||| |
Db        237 ADCGLNVHKQCSKMVPNDCKPDLKHVK-KVYSCDLTTLVKAHTTKRPMVVDMCIREIESR 295

Qy        377 GLTETGLYRISGCDRTVKELKEKFLR--VKTVPLLSKVDDIHAICSLLKDFLRNLKEPLL 434
              ||   |||:||    ::::|  | |   |    ::  :||: |   || : |:|  ||:
Db        296 GLNSEGLYRVSGFSDLIEDVKMAFDRDGEKADISVNMYEDINIITGALKLYFRDLPIPLI 355

Qy        435 TFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTK---M 491
              |:    |:|:|:| | |  :  :::|:  || |: :|| :|| ||:||  : | |   |
Db        356 TYDAYPKFIESAKIMDPDEQLETLHEALKLLPPAHCETLRYLMAHLKRV--TLHEKENLM 413

Qy        492 DVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLL 532
              :  ||  |||||::   |  | :    ||: |  ||| |:
Db        414 NAENLGIVFGPTLMRS--PELDAMAALNDIRYQRLVVELLI 452


RESULT 14
Q92153
ID   Q92153      PRELIMINARY;      PRT;   299 AA.
AC   Q92153;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HAT-2.
GN   HAT-2.
OS   Serinus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;
OC   Fringillidae; Carduelinae; Serinus.
OX   NCBI_TaxID=54072;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92252621; PubMed=1374499;
RA   George J.M., Clayton D.F.;
RT   "Differential regulation in the avian song control circuit of an mRNA
RT   predicting a highly conserved protein related to protein kinase C and
RT   the bcr oncogene.";
RL   Brain Res. Mol. Brain Res. 12:323-329(1992).
DR   EMBL; S98891; AAB22136.1; -.
DR   HSSP; P28867; 1PTQ.
DR   InterPro; IPR002219; DAG_PE-bind.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   PRINTS; PR00008; DAGPEDOMAIN.
DR   SMART; SM00109; C1; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
SQ   SEQUENCE   299 AA;  34334 MW;  8FAAD0A7EC6AFAE4 CRC64;

  Query Match           9.9%;  Score 320;  DB 13;  Length 299;
  Best Local Similarity   31.6%;  Pred. No. 4.3e-14;
  Matches   92;  Conservative   47;  Mismatches  106;  Indels   46;  Gaps   10;
```

```
Qy        259 STLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLS--LKC 316
Db         31 ATLKENEHVPKYE-------------KVHNFKVHTFRGPHWCEYCA-NFMWGLIAQGVKC 76

Qy        317 RDCRVVSHPECRDRCPLPCIPTL----------IGTPVKIGEGMLADFVSQTSPMIPSIV 366
Db         77 ADCGLNVHKQCSKMVPNDCKPDLKHVKKVYSCDLTTLVK------AHFTKR-----PMVV 125

Qy        367 VHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLR--VKTVPLLSKVDDIHAICSLLKD 424
Db        126 DMCIREIESRGLNSEGLYRVSGFSDLIEDVKMAFDRDGEKADISVNMYEDINIITGALKL 185

Qy        425 FLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVA 484
Db        186 YFRDLPIPLITYDAYPKFIESAKTTDPDEQLEILHEALKLLPPAHCETLRYLMAHLKRV- 244

Qy        485 QSPHTK---MDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLL 532
Db        245 -TLHEKENLMSAENLGIVFGPTLM--RAPELDAMAALNDIRYQRLVVEMLI 292


RESULT 15
O94466
ID   O94466      PRELIMINARY;      PRT;   695 AA.
AC   O94466;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Putative GTPase-activating protein C23G7.08C.
GN   SPBC23G7.08C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Xiang Z., Aves S., Wood V., Rajandream M.A., Barrell B.G.;
RL   Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO OTHER EUKARYOTIC RHO GDP-DISSOCIATION INHIBITORS.
DR   EMBL; AL035065; CAA22624.1; -.
DR   HSSP; Q07960; 1RGP.
DR   InterPro; IPR001060; Cdc15_Fes_CIP4.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00611; FCH; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00055; FCH; 1.
DR   SMART; SM00324; RhoGAP; 1.
KW   Hypothetical protein; GTPase activation.
FT   DOMAIN      420    426       POLY-SER.
SQ   SEQUENCE   695 AA;  76678 MW;  1DEAF1328D576263 CRC64;

  Query Match             9.5%;  Score 308;  DB 3;  Length 695;
  Best Local Similarity   23.7%;  Pred. No. 9.2e-13;
  Matches  137;  Conservative   87;  Mismatches  223;  Indels  132;  Gaps   23;

Qy         35 KDFEDFRKKWQRTDHE----LGKYKDLLMKAETE-RSALDVKLKHARNQVDVEIK---RR 86
Db        145 KSIKEYAKKQENAYLEAVMQMDKSKSRFKGAETEYNRALDNK-----NTGDSQKKVGFFK 199

Qy         87 QRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE-QKSALAF----------LNRGQPS 134
Db        200 PKSNAQLTKLEDEARLKAENAESDMHSKIENAQNVQKQLLCIHRPNYIKQFFSLQREIES 259

Qy        135 SSNAGNKRLSTIDESGSILSDI---------------SFDKTDESLDWDSSLVKTFKLK 178
Db        260 SLIANYLRYTKLCESNTLLNGLTIRPQKPTPTNCGLQHALDNINANTDFVQYVLHA-SIK 318

Qy        179 KREKRRST--SRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIE----A 232
Db        319 HEDNKNPTDASKTKIIQPPSSYGTGSSAGKTNPPVNPTI--KVTAAIPS---PLQNTNPA 373

Qy        233 VSTIETVPYWTRSRRKTGTLQPWNSDSTLN--SRQLEPRTETDSVGTPQSNGGMRLHDFV 290
Db        374 PSTFPNPSVASPAFPNSSTSNPSTAPASASPLASTLKPSTANDTNGSSSSSSS------- 426

Qy        291 SKTVIKPESCVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCP---LPCIP----TLIGT- 342
Db        427 -----NPRTSSPLASNAE-------------NKPPVAQQSPPVLLPTLPPIQTTTIQTS 467

Qy        343 ------PVKIGEGMLADFVSQTS-------------------------PMIPSIVVHCV 370
Db        468 REVAPPPSSINSNRAASPFRPTSVSPQPSSPTKSLLFGARLDAIILREHSNIPNIVMQCT 527

Qy        371 NEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLL------SKVDDIHAICSLLKD 424
Db        528 SQVENFGLNLQGIYRVPSSSARVNMLRSQF---ENNPLLQLHTPEDYENDVHAVADLLKI 584

Qy        425 FLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVA 484
Db        585 FFRELREPLIPDNHQRDFIDAGNVEDESRRRDAVHRAINDLPDANYSTIRHLTIHLAKIK 644

Qy        485 Q-SPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIK 522
Db        645 ENSDVNKMSTNNLAIIWGPTIIKQAT-IPEISSFSRTIE 682


Search completed: July  3, 2003, 09:56:17
Job time : 92 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 3, 2003, 09:44:06 ; Search time 73 Seconds
(without alignments)
1153.622 Million cell updates/sec

Title: US-09-881-736-2
Perfect score: 3243
Sequence: 1 MDTMMLNVRNLFEQLVRRVE.....SKSATNLGRQGNFFASPMLK 632

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3233	99.7	632	22	AAM40047	Human polypeptide
2	3230	99.6	632	23	AAU69419	Lung small cell ca
3	2919	90.0	570	23	AAU69422	Lung small cell ca
4	1182	36.4	255	22	AAB94223	Human protein sequ
5	971.5	30.0	256	22	AAU32069	Novel human secret
6	864	26.6	625	22	ABB68056	Drosophila melanog
7	625	19.3	4318	22	ABG14787	Novel human diagno
8	618	19.1	665	22	ABG10230	Novel human diagno
9	618	19.1	1086	22	ABG19904	Novel human diagno
10	618	19.1	1139	22	ABG13958	Novel human diagno
11	581	17.9	119	22	AAM41833	Human polypeptide
12	477	14.7	384	22	ABB59247	Drosophila melanog
13	322	9.9	334	22	AAM39659	Human polypeptide
14	322	9.9	352	22	AAM41445	Human polypeptide
15	301	9.3	1261	19	AAW75995	GTPase activating
16	301	9.3	1261	21	AAY90268	Human GTP-ase acti
17	294.5	9.1	193	22	AAG67553	Amino acid sequenc
18	290.5	9.0	736	20	AAW97809	Human GTPase regul
19	290.5	9.0	759	22	ABG16793	Novel human diagno
20	290.5	9.0	800	22	ABG16794	Novel human diagno
21	286.5	8.8	856	21	AAB41660	Human ORFX ORF1424
22	286.5	8.8	1173	22	AAB97911	Human G-protein ac
23	278.5	8.6	967	22	ABG12188	Novel human diagno
24	270.5	8.3	589	7	AAP60303	Sequence encoded b
25	270.5	8.3	2548	20	AAY05781	Human myosin IXa.
26	267	8.2	1212	22	ABG20843	Novel human diagno
27	265	8.2	494	22	ABB59717	Drosophila melanog
28	263.5	8.1	555	22	AAU17313	Novel signal trans
29	263	8.1	985	22	ABG12190	Novel human diagno
30	260.5	8.0	802	20	AAY21700	Human oligophrenin
31	258	8.0	903	21	AAB42926	Human ORFX ORF2690
32	256.5	7.9	659	23	ABP51281	Human MDDT SEQ ID
33	251	7.7	170	22	AAG65821	RhoGAP domain from
34	251	7.7	170	22	AAG67552	Amino acid sequenc
35	249	7.7	2110	22	ABB58077	Drosophila melanog
36	248.5	7.7	526	22	AAB93466	Human protein sequ
37	247.5	7.6	291	22	AAU17333	Novel signal trans
38	246.5	7.6	296	22	AAM25391	Human protein sequ
39	246.5	7.6	718	22	ABB66573	Drosophila melanog
40	246	7.6	816	22	AAG66505	GTP enzyme Rho fam
41	241.5	7.4	1844	22	ABB58723	Drosophila melanog
42	238	7.3	974	23	ABG61872	Prostate cancer-as
43	232	7.2	322	22	ABG06669	Novel human diagno
44	231	7.1	818	23	ABB90757	Human Tumour Endot
45	229.5	7.1	668	22	ABG27731	Novel human diagno

ALIGNMENTS

RESULT 1
AAM40047
ID AAM40047 standard; Protein; 632 AA.
XX
AC AAM40047;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 3192.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.

XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI59203.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 4; SEQ ID NO 3192; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 632 AA;

Query Match 99.7%; Score 3233; DB 22; Length 632;
Best Local Similarity 99.7%; Pred. No. 6.8e-263;
Matches 630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60

Qy         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120

Qy        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180

Qy        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240

Qy        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300

Qy        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360

Qy        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420

Qy        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480

Qy        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMLQDIKRQPKVVERLLSLPLEYWS 540
```

```
Qy        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSGSLSQRVRS 600

Qy        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              ||||||||||||||||||||||||||||||||
Db        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
```

RESULT 2
AAU69419
ID AAU69419 standard; Protein; 632 AA.
XX
AC AAU69419;
XX
DT 30-JAN-2002 (first entry)
XX
DE Lung small cell carcinoma antigen #13.
XX
KW Human; cytostatic; antitumour; lung small cell cancer antigen;
KW tumour; lung cancer.
XX
OS Homo sapiens.
XX
PN WO200177168-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001WO-US11859.
XX
PR 11-APR-2000; 2000US-196780P.
PR 21-JUN-2000; 2000US-213361P.
PR 01-SEP-2000; 2000US-229763P.
PR 05-SEP-2000; 2000US-230629P.
PR 14-SEP-2000; 2000US-232565P.
PR 19-DEC-2000; 2000US-257037P.
PR 08-JAN-2001; 2001US-260796P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Lodes MJ, Wang T, Mohamath R, Indirias CY;
XX
DR WPI; 2002-010896/01.
DR N-PSDB; AAS61859.
XX
PT Lung tumour polynucleotide and polypeptides useful in therapy and
PT diagnosis of cancer especially lung cancer -
XX
PS Claim 2; Page 270-271; 295pp; English.
XX
CC The invention relates to novel isolated lung small cell cancer antigen
CC polynucleotides (I) and polypeptides (II) used in a method of detecting
CC cancer in a patient. The method is optionally performed by
CC utilising oligonucleotides (III), where the biological sample
CC from the patient is contacted with (III), detecting the amount of
CC polynucleotide hybridised to (III) in the sample and comparing the
CC amount of polynucleotide to a predetermined cut-off value and thereby
CC determining cancer in a patient. (I), (II) or antigen-presenting cells
CC expressing (II) is useful for stimulating and/or expanding T cells
CC specific for a tumour protein. The method comprises contacting T cells
CC with one of the components under conditions to permit the stimulation
CC and/or expansion of the cells. A composition comprising (I) is useful for
CC stimulating an immune response in a patient and for inhibiting the
CC development of a cancer especially lung cancer in a patient. An
CC isolated T cell population is useful for removing tumour cells from the
CC biological sample and for inhibiting the development of cancer in a
CC patient. AAU69407-AAU69431 represent novel human lung small cell
CC cancer antigen amino acid sequences of the invention.
XX
SQ Sequence 632 AA;

Query Match 99.6%; Score 3230; DB 23; Length 632;

```
  Best Local Similarity   99.7%;  Pred. No. 1.2e-262;
  Matches  630;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60

Qy         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120

Qy        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSHISFDKTDESLDWDSSLVKTFKLKKR 180

Qy        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240

Qy        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300

Qy        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360

Qy        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420

Qy        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480

Qy        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMLQDIKRQPKVVERLLSLPLEYWS 540

Qy        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600

Qy        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              ||||||||||||||||||||||||||||||||
Db        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
```

RESULT 3
AAU69422

```
ID   AAU69422 standard; Protein; 570 AA.
XX
AC   AAU69422;
XX
DT   30-JAN-2002  (first entry)
XX
DE   Lung small cell carcinoma antigen #16.
XX
KW   Human; cytostatic; antitumour; lung small cell cancer antigen;
KW   tumour; lung cancer.
XX
OS   Homo sapiens.
XX
PN   WO200177168-A2.
XX
PD   18-OCT-2001.
XX
PF   11-APR-2001; 2001WO-US11859.
XX
PR   11-APR-2000; 2000US-196780P.
PR   21-JUN-2000; 2000US-213361P.
PR   01-SEP-2000; 2000US-229763P.
PR   05-SEP-2000; 2000US-230629P.
PR   14-SEP-2000; 2000US-232565P.
PR   19-DEC-2000; 2000US-257037P.
PR   08-JAN-2001; 2001US-260796P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Lodes MJ,  Wang T,  Mohamath R,  Indirias CY;
XX
DR   WPI; 2002-010896/01.
DR   N-PSDB; AAS61862.
XX
PT   Lung tumour polynucleotide and polypeptides useful in therapy and
PT   diagnosis of cancer especially lung cancer  -
XX
PS   Claim 2; Page 282-284; 295pp; English.
XX
CC   The invention relates to novel isolated lung small cell cancer antigen
CC   polynucleotides (I) and polypeptides (II) used in a method of detecting
CC   cancer in a patient. The method is optionally performed by
CC   utilising oligonucleotides (III), where the biological sample
CC   from the patient is contacted with (III), detecting the amount of
CC   polynucleotide hybridised to (III) in the sample and comparing the
CC   amount of polynucleotide to a predetermined cut-off value and thereby
CC   determining cancer in a patient. (I), (II) or antigen-presenting cells
CC   expressing (II) is useful for stimulating and/or expanding T cells
CC   specific for a tumour protein. The method comprises contacting T cells
CC   with one of the components under conditions to permit the stimulation
CC   and/or expansion of the cells. A composition comprising (I) is useful for
CC   stimulating an immune response in a patient and for inhibiting the
CC   development of a cancer especially lung cancer in a patient. An
CC   isolated T cell population is useful for removing tumour cells from the
CC   biological sample and for inhibiting the development of cancer in a
CC   patient. AAU69407-AAU69431 represent novel human lung small cell
CC   cancer antigen amino acid sequences of the invention.
XX
SQ   Sequence   570 AA;

  Query Match             90.0%;  Score 2919;  DB 23;  Length 570;
  Best Local Similarity   99.8%;  Pred. No. 1.5e-236;
  Matches  569;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         63 TERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEEQK 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEEQK 60

Qy        123 SALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKREK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKREK 120

Qy        183 RRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVPYW 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVPYW 180

Qy        243 TRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVP 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVP 240

Qy        303 CGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSPMI 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSPMI 300

Qy        363 PSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICSLL 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 PSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICSLL 360

Qy        423 KDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQR 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQR 420

Qy        483 VAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWSQF 542
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        421 VAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMLQDIKRQPKVVERLLSLPLEYWSQF 480
```

```
Qy        543 MMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRSTL 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 MMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRSTL 540

Qy        603 TKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              ||||||||||||||||||||||||||||||
Db        541 TKNTPRFGSKSKSATNLGRQGNFFASPMLK 570


RESULT 4
AAB94223
ID   AAB94223 standard; Protein; 255 AA.
XX
AC   AAB94223;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human protein sequence SEQ ID NO:14587.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 8; SEQ ID 14587; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence   255 AA;

  Query Match            36.4%;  Score 1182;  DB 22;  Length 255;
  Best Local Similarity  99.2%;  Pred. No. 6e-91;
  Matches  234;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        397 KEKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIA 456
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 KLKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIA 79

Qy        457 AMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVT 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 AMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVT 139

Qy        517 MSQDIKRQPKVVERLLSLPLEYWSQFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLG 576
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 MLQDIKRQPKVVERLLSLPLEYWSQFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLG 199

Qy        577 PVTTPEHQLLKTPSSSSLSQRVRSTLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 PVTTPEHQLLKTPSSSSLSQRVRSTLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 255


RESULT 5
AAU32069
ID   AAU32069 standard; Protein; 256 AA.
XX
AC   AAU32069;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Novel human secreted protein #2560.
XX
KW   Human; vaccination; gene therapy; nutritional supplement;
KW   stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW   immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200179449-A2.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; 2001WO-US08656.
XX
PR   18-APR-2000; 2000US-0552929.
PR   26-JAN-2001; 2001US-0770160.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-611725/70.
XX
PT   Nucleic acids encoding a range of human polypeptides, useful in genetic
PT   vaccination, testing and therapy -
XX
PS   Claim 20; Page 551; 765pp; English.
XX
CC   The invention relates to novel human secreted polypeptides. The
CC   polypeptides and antibodies to the polypeptides are useful for
CC   determining the presence of or predisposition to a disease associated
CC   with altered levels of polypeptide. The polypeptides are also useful for
CC   identifying agents (agonists and antagonists) that bind to them. Cells
CC   expressing the proteins are useful for identifying a therapeutic agent
CC   for use in treatment of a pathology related to aberrant expression or
CC   physiological interactions of the polypeptide. Vectors comprising
CC   the nucleic acids encoding the polypeptides and cells genetically
CC   engineered to express them are also useful for producing the proteins.
CC   The proteins are useful in genetic vaccination, testing and
CC   therapy, and can be used as nutritional supplements. They may be used to
CC   increase stem cell proliferation; to regulate haematopoiesis; and in
CC   bone, cartilage, tendon  and/or nerve tissue growth or regeneration;
CC   immune suppression and/or stimulation; as anti-inflammatory agents; and
```

```
CC   in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC   sequences of novel human secreted proteins of the invention.
XX
SQ   Sequence   256 AA;

  Query Match            30.0%;  Score 971.5;  DB 22;  Length 256;
  Best Local Similarity   84.0%;  Pred. No. 3.1e-73;
  Matches  205;  Conservative    3;  Mismatches   27;  Indels    9;  Gaps    4;

Qy        357 QTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIH 416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 QTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIH 60

Qy        417 AICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFL 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFL 120

Qy        477 MIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKR-QPKVVERLLSLP 535
              |||||||||||||||||||||||||||||||||||| |  :  |||||||| |
Db        121 MIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMLQGHQGVQPKVVERLLFLG 180

Qy        536 L-EYWSQFMMVEQE---NIDPLHVIENSNAFSTPQTPDIKV----SLLGPVTTPEHQLLK 587
              | |||    |   |     ||||||||||||||||||      |    |  | :: |
Db        181 LWEYWEFSFMDGWEQGGTFDPLHVIENSNAFSTPQTPDIKAVPGGGLCVHFTAGEAEIQK 240

Qy        588 TPSS 591
                | |
Db        241 GPPS 244


RESULT 6
ABB68056
ID   ABB68056 standard; Protein; 625 AA.
XX
AC   ABB68056;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 30960.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL12159.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 30960; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   625 AA;

  Query Match            26.6%;  Score 864;  DB 22;  Length 625;
  Best Local Similarity   35.1%;  Pred. No. 1.5e-63;
  Matches  228;  Conservative  110;  Mismatches  225;  Indels   86;  Gaps   21;

Qy          4 MMLNVRNLFEQLVRRVEILSEGN-EVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMKAE 62
              | |:    |: | | :::|::|  | :|::  : || : :|      |  : :: | |:
Db          1 MALSALASFDDLRRCMQVLTDGTPEEEFLRFLRMFEQYHEKCAGYAAETARIQNELDKSL 60

Qy         63 TERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEEQK 122
              |:   |: || |||  :|:||| |::|| : : :| :|  : ::|  : :   |: | :
Db         61 TKMGDLEGKLFHARRIIDMEIKARRQAEHERDAMESKIMAVADLLRHERN----LNNETR 116

Qy        123 SALAFLNRGQPSSS-----NA--GNKRLSTIDESGSILSDISFDKT-DESLDWDSSLVKT 174
                ||||:   |||     ||   :|    |: :||:|||:|   : |: ||     |:|
Db        117 DKLAFLHT-LPSSRKRKSLNAVREDKSYGDINSTGSLLSDLSITHSEDDFLD-----VRT 170

Qy        175 FKLKKREKRRSTSRQFV-----------------DGPPGPVKKTR--SIGSAVD----- 209
              |   || | |  :  :                  |      |:|  |:|  |:
Db        171 SK-SWREHRPSLPKNQIPSVGNKRSRLSTGLNGSMSGTTPTTGKSRRSSVGIGVEQHTVD 229

Qy        210 --QGNESIVAKTTVTVPNDG-GPIEAVSTIETVPYWTRSRRKTGTLQPWNSDSTLNSRQL 266
                || |   | | ||:| || | | | ||||::|    : |    |      | |
Db        230 VGQGAERFCATTKVTIPQDGQGVIRAESTIESLPVIAGNERIGDGLSSTPRRSVLKEATA 289

Qy        267 EPRTETDSV--------GTP-QSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLSLKCR 317
              | |  :::        ||| |    || | |  || :: ::|| | |||:|| : |:||
Db        290 PPLTPVNAMAPHVVAESGTPLQHRPLMRNHTFSQKTFLRGDNCVQCQKRIRFGAVGLRCR 349

Qy        318 DCRVVSHPECRDRCPLPCIPTLIGTP-VKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQR 376
              || |  | :||    : |:|   |||  |   | : ||    :||||:::|||||||| |
Db        350 DCPVRCHIDCRYLLTVSCVPQ-TGTPTTKTMTGYVTDFAPSIAPMIPALIVHCVNEIEAR 408

Qy        377 GLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTF 436
              |||| ||||:|  :|  | |||:||| |  | |    ||: :|  :|||||:| |||:
Db        409 GLTEVGLYRLSSSEREYKALKEQFLRGKATPHLGNT-DIYVLCCCVKDFLRSLTEPLIPT 467

Qy        437 RLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTKMDVANL 496
                : |  | :  |   :  : ::| :|||||||||||::| ||:|| |   | : |:
Db        468 SQWKDFANAVQNPDTKTAQDMLVKSVKQLPQANRDTLAFLILHFQRIAQCPVVLMPIDNI 527

Qy        497 AKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWSQFMMVEQENIDPLHVI 556
              : :|||||| :: |:||   :  ::  | :|:: || ||: :| |:::     |||
Db        528 SLIFGPTIVGYSTPDPDQHAIYTEVFTQKQVMKALLELPVSFWEQYIV-----IDPTR-- 580

Qy        557 ENSNAFSTPQT-----PDIKVSLLGPVTTP-------EHQLLKTPSSSS 593
                    || |     |  |  ||    ||       : :   || :|:
Db        581 -------TPATVIKRVPSNKNDLLSLYATPFKGGTIKKRKFYGTPPASA 622


RESULT 7
ABG14787
ID   ABG14787 standard; Protein; 4318 AA.
XX
AC   ABG14787;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #14778.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
```

```
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS78974.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 20; SEQ ID No 45146; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   4318 AA;

  Query Match             19.3%;  Score 625;  DB 22;  Length 4318;
  Best Local Similarity   38.8%;  Pred. No. 4.3e-42;
  Matches  167;  Conservative   46;  Mismatches   85;  Indels  132;  Gaps   16;

Qy        216 VAKTTVTVPNDGGPIE--------AVSTIETVPYWTRSR-----RKTGTLQ--------- 253
              :||: ::  |  | |           :  :|  || : |      :||  :
Db       3915 IAKSILSQKNKAGGITLPDFKLYYKATVTKTTWYWYQHRAVDQWNRTGPSEIMPHIYNYL 3974

Qy        254 ---------PWNSDSTLNS----------RQLE------PRTETDS------------VG 276
                      |  ||  |             |:|:      | |: :|            :
Db       3975 IFDKPDKNKQWGKDSLFNKWCWENWLAICRKLKLNPFLTPYTKINSRWIKDLHVRPKTIK 4034

Qy        277 TPQSNGGMRLH------DFVSKTVIKPESCVPCGKRIKFGKLSLKCRDCRVVSHPECRDR 330
              | : | |  :       ||:|||   |::     |  |:  : ||   |
Db       4035 TLEENLGNTIQVIGMGKDFMSKT---PKAMATKAKIDKWDLIKLKSFFC----------- 4080

Qy        331 CPLPCIPTLIGTPVKIG------EGMLADFVSQTSPMIPSIVVHCVNEIEQ--------- 375
                      |   | :::       | : | : |          ::    ||::|
Db       4081 -------TAKETTIRVNRQPTEWEKIFATYSSD-----KGLISRIYNELKQIYKKKTNNP 4128

Qy        376 ---------RGLTETGLY----RISGCDR--TVKELKEK-FLRVKTVPLL---------- 409
                      |  ::  :|     :  |     ::|:: |  :|    |:
Db       4129 IKKWTNDMNRHFSKEDIYAAKKHMKKCSSLPAIREMQIKTTMRYHLTPVRIAIIKKSGNN 4188

Qy        410 SKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQAN 469
              ||||||||||||||||||| ||||||||||:|||||||||||||||:||||||||||||
Db       4189 SKVDDIHAICSLLKDFLRNFKEPLLTFRLNKAFMEAAEITDEDNSISAMYQAVGELPQAN 4248

Qy        470 RDTLAFLMIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVE 529
              |||| |||||||||||||:|||:|||||:||| ||||||||||:|||| |||| ||||||
Db       4249 RDTLVFLMIHLQRVAQSPYTKMNVANLAEVFGSTIVAHAVPNPEPVTMLQDIKCQPKVVE 4308

Qy        530 RLLSLPLEYW 539
              || |||||||
Db       4309 RLPSLPLEYW 4318


RESULT 8
ABG10230
ID   ABG10230 standard; Protein; 665 AA.
XX
AC   ABG10230;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #10221.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS74417.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 20; SEQ ID No 40589; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   665 AA;

  Query Match             19.1%;  Score 618;  DB 22;  Length 665;
```

```
  Best Local Similarity   46.9%;  Pred. No. 8.1e-43;
  Matches  153;  Conservative   35;  Mismatches   68;  Indels   70;  Gaps   11;

Qy        255 WNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLSL 314
              |  |  || |    :|  :::|    : ||   ||:|||   |::     |  |:  : |
Db        369 WIKD--LNVRPKTIKTLEENLGNTIQDIGMG-KDFMSKT---PKAMATKTKIDKWDLIQL 422

Qy        315 KCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIG------EGMLADFVSQTSPMIPSIVVH 368
              |   |                  |   | :::       | : | : |        ::
Db        423 KSFFC-----------------TAKETTIRVNRQPTEWEKIFATYSSD-----KGLISR 459

Qy        369 CVNEIEQ-----------------RGLTETGLY----RISGCDR--TVKELKEK-FLRV 403
                ||::|                 |  ::  :|     :  |     ::|:: |  :|
Db        460 IYNELKQIYKKKTNNPIKKWTNDMNRHFSKEDIYAAKKHMKKCSSLPAIREMQIKTTMRY 519

Qy        404 KTVPLL----------SKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDN 453
                |:           ||||||||||||||||||| ||||||||||:||||||||||||
Db        520 HLTPVRIAIIKKSGNNSKVDDIHAICSLLKDFLRNFKEPLLTFRLNKAFMEAAEITDEDN 579

Qy        454 SIAAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPD 513
              ||:||||||||||||||| ||||||||||||:|||:|||||:||| ||||||||||:
Db        580 SISAMYQAVGELPQANRDTLVFLMIHLQRVAQSPYTKMNVANLAEVFGSTIVAHAVPNPE 639

Qy        514 PVTMSQDIKRQPKVVERLLSLPLEYW 539
              |||| |||| |||||||| |||||||
Db        640 PVTMLQDIKCQPKVVERLPSLPLEYW 665


RESULT 9
ABG19904
ID   ABG19904 standard; Protein; 1086 AA.
XX
AC   ABG19904;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #19895.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS84091.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 20; SEQ ID No 50263; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC  .and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   1086 AA;

  Query Match             19.1%;  Score 618;  DB 22;  Length 1086;
  Best Local Similarity   46.9%;  Pred. No. 1.8e-42;
  Matches  153;  Conservative   35;  Mismatches   68;  Indels   70;  Gaps   11;

Qy        255 WNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLSL 314
              |  |  || |    :|  :::|    : ||   ||:|||   |::     |  |:  : |
Db        790 WIKD--LNVRPKTIKTLEENLGNTIQDIGMG-KDFMSKT---PKAMATKTKIDKWDLIQL 843

Qy        315 KCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIG------EGMLADFVSQTSPMIPSIVVH 368
              |   |                  |   | :::       | : | : |        ::
Db        844 KSFFC-----------------TAKETTIRVNRQPTEWEKIFATYSSD-----KGLISR 880

Qy        369 CVNEIEQ-----------------RGLTETGLY----RISGCDR--TVKELKEK-FLRV 403
                ||::|                 |  ::  :|     :  |     ::|:: |  :|
Db        881 IYNELKQIYKKKTNNPIKKWTNDMNRHFSKEDIYAAKKHMKKCSSLPAIREMQIKTTMRY 940

Qy        404 KTVPLL----------SKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDN 453
                |:           ||||||||||||||||||| ||||||||||:||||||||||||
Db        941 HLTPVRIAIIKKSGNNSKVDDIHAICSLLKDFLRNFKEPLLTFRLNKAFMEAAEITDEDN 1000

Qy        454 SIAAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPD 513
              ||:||||||||||||||| ||||||||||||:|||:|||||:||| ||||||||||:
Db       1001 SISAMYQAVGELPQANRDTLVFLMIHLQRVAQSPYTKMNVANLAEVFGSTIVAHAVPNPE 1060

Qy        514 PVTMSQDIKRQPKVVERLLSLPLEYW 539
              |||| |||| |||||||| |||||||
Db       1061 PVTMLQDIKCQPKVVERLPSLPLEYW 1086


RESULT 10
ABG13958
ID   ABG13958 standard; Protein; 1139 AA.
XX
AC   ABG13958;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #13949.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
```

```
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS78145.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 20; SEQ ID No 44317; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   1139 AA;

  Query Match           19.1%;  Score 618;  DB 22;  Length 1139;
  Best Local Similarity  45.0%;  Pred. No. 1.9e-42;
  Matches  158;  Conservative   39;  Mismatches   70;  Indels   84;  Gaps   14;

Qy       242 WTRSRRKTGTLQPWNSDST-LNSR-----QLEPRTETDSVGTPQSNGGMRLH------DF 289
             |    ||   | |: :  | :|||      : |:|    : | : | |: :       ||
Db       820 WLAICRKL-KLDPFLTAYTKINSRWIKDLNIRPKT----IKTLEENLGITIQDIGMGKDF 874

Qy       290 VSKTVIKPESCVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIG-- 347
             :|||   |::     |  |:  : ||   |                   |   | :::
Db       875 MSKT---PKAMATKTKIDKWDLIQLKSFFC-----------------TAKETTIRVNRQ 913

Qy       348 ----EGMLADFVSQTSPMIPSIVVHCVNEIEQ-----------------RGLTETGLY- 384
                 | : | : |        ::    ||::|                 |  ::  :|
Db       914 PTEWEKIFATYSSD-----KGLISRIYNELKQIYKKKTNNPIKKWTNDMNRHFSKEDIYA 968

Qy       385 ---RISGCDR--TVKELKEK-FLRVKTVPLL----------SKVDDIHAICSLLKDFLRN 428
                 :  |     ::|:: |  :|   |:           ||||||||||||||||||
Db       969 AKKHMKKCSSLPAIREMQIKTTMRYHLTPVRIAIIKKSGNNSKVDDIHAICSLLKDFLRN 1028

Qy       429 LKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQSPH 488
             ||||||||||:|||||||||||||||:|||||||||||||||| ||||||||||||:
Db      1029 FKEPLLTFRLNKAFMEAAEITDEDNSISAMYQAVGELPQANRDTLVFLMIHLQRVAQSPY 1088

Qy       489 TKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYW 539
             |||:|||||:||| |||||||||:|||| |||| |||||||| |||||||
Db      1089 TKMNVANLAEVFGSTIVAHAVPNPEPVTMLQDIKCQPKVVERLPSLPLEYW 1139


RESULT 11
AAM41833
ID   AAM41833 standard; Protein; 119 AA.
XX
AC   AAM41833;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 6764.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI60989.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 2; SEQ ID NO 6764; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   119 AA;

  Query Match           17.9%;  Score 581;  DB 22;  Length 119;
  Best Local Similarity  99.1%;  Pred. No. 6.5e-41;
  Matches  115;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60
             ||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4 MDTMMLSVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 63

Qy        61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQ 116
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        64 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQ 119


RESULT 12
ABB59247
```

```
ID   ABB59247 standard; Protein; 384 AA.
XX
AC   ABB59247;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 4533.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL03350.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 4533; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   384 AA;

  Query Match             14.7%;  Score 477;  DB 22;  Length 384;
  Best Local Similarity   40.7%;  Pred. No. 2.4e-31;
  Matches  110;  Conservative   44;  Mismatches  102;  Indels   14;  Gaps    6;

Qy        279 QSNGG-MRLHDFVSKT-VIKPESCVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCI 336
              ||: | :| |:|  |:      :|| | |||:|   ||:|| | :  |  |  :  : ||
Db         78 QSHSGLLREHNFKIKSYYYNVGNCVHCRKRIRFAMASLRCRACPLRCHIGCCRQLTVNCI 137

Qy        337 P-TLIGTPVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKE 395
              |   |||      | |:|:  : :||:|:::|||| ||| ||| : ||||:|      |
Db        138 PQPQIGTK----RGCLSDYAPRVAPMVPALIVHCVTEIEARGLQQEGLYRVSSTREKCKR 193

Qy        396 LKEKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSI 455
              |: | || |: | |    | | :|  :|||| |  ||:     | | ||    |
Db        194 LRRKLLRGKSTPHLGN-KDTHTLCCCVKDFLRQLVHPLIPIYHRRDFEEATRHEDRLAVE 252

Qy        456 AAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPV 515
               |:| || || ||:|||||:||:| |::|:||  :| | ||| :| ||:        |
Db        253 MAVYLAVLELHQAHRDTLAYLMLHWQKIAESPAVRMTVNNLAVIFAPTLFG------DLD 306

Qy        516 TMSQDIKRQPKVVERLLSLPLEYWSQFMMV 545
                :::    :|:: || :|  :||||: |
Db        307 LTLENVVTWQRVLKVLLLMPAGFWSQFLEV 336
```

```
RESULT 13
AAM39659
ID   AAM39659 standard; Protein; 334 AA.
XX
AC   AAM39659;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 2804.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI58815.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 4; SEQ ID NO 2804; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   334 AA;

  Query Match             9.9%;  Score 322;  DB 22;  Length 334;
  Best Local Similarity   32.0%;  Pred. No. 2.2e-18;
  Matches   90;  Conservative   48;  Mismatches  117;  Indels   26;  Gaps    8;

Qy        259 STLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLS--LKC 316
```

```
              :||   :  |: |            ::|:|   |   |  |  |     :| ::  :||
Db         66 ATLKENEQIPKYE------------KIHNFKVHTFRGPHWCEYCA-NFMWGLIAQGVKC 111

Qy        317 RDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQR 376
              || :  | :|    |  | | |     |:    |   |   :  | :|  |: ||| |
Db        112 ADCGLNVHKQCSKMVPNDCKPDLKHVK-KVYSCDLTTLVKAHTTKRPMVVDMCIREIESR 170

Qy        377 GLTETGLYRISGCDRTVKELKEKFLR--VKTVPLLSKVDDIHAICSLLKDFLRNLKEPLL 434
              ||   ||||:||    ::::|  | |   |    ::  :||: |   || : |:|  ||:
Db        171 GLNSEGLYRVSGFSDLIEDVKMAFDRDGEKADISVNMYEDINIITGALKLYFRDLPIPLI 230

Qy        435 TFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTK---M 491
              |:     |:|:|:| | |  :  :::|:  || |: :|| :|| ||:||  : | |   |
Db        231 TYDAYPKFIESAKIMDPDEQLETLHEALKLLPPAHCETLRYLMAHLKRV--TLHEKENLM 288

Qy        492 DVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLL 532
              :  ||  |||||::   |  | :   ||: |  ||| |:
Db        289 NAENLGIVFGPTLMRS--PELDAMAALNDIRYQRLVVELLI 327


RESULT 14
AAM41445
ID   AAM41445 standard; Protein; 352 AA.
XX
AC   AAM41445;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 6376.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI60601.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 2; SEQ ID NO 6376; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   352 AA;

  Query Match             9.9%;  Score 322;  DB 22;  Length 352;
  Best Local Similarity   32.0%;  Pred. No. 2.3e-18;
  Matches   90;  Conservative   48;  Mismatches  117;  Indels   26;  Gaps    8;

Qy        259 STLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLS--LKC 316
              :||   :  |: |            ::|:|   |   |  |  |     :| ::  :||
Db         84 ATLKENEQIPKYE------------KIHNFKVHTFRGPHWCEYCA-NFMWGLIAQGVKC 129

Qy        317 RDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQR 376
              || :  | :|    |  | | |     |:    |   |   :  | :|  |: ||| |
Db        130 ADCGLNVHKQCSKMVPNDCKPDLKHVK-KVYSCDLTTLVKAHTTKRPMVVDMCIREIESR 188

Qy        377 GLTETGLYRISGCDRTVKELKEKFLR--VKTVPLLSKVDDIHAICSLLKDFLRNLKEPLL 434
              ||   ||||:||    ::::|  | |   |    ::  :||: |   || : |:|  ||:
Db        189 GLNSEGLYRVSGFSDLIEDVKMAFDRDGEKADISVNMYEDINIITGALKLYFRDLPIPLI 248

Qy        435 TFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTK---M 491
              |:     |:|:|:| | |  :  :::|:  || |: :|| :|| ||:||  : | |   |
Db        249 TYDAYPKFIESAKIMDPDEQLETLHEALKLLPPAHCETLRYLMAHLKRV--TLHEKENLM 306

Qy        492 DVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLL 532
              :  ||  |||||::    |  | :    ||: |  ||| |:
Db        307 NAENLGIVFGPTLMRS--PELDAMAALNDIRYQRLVVELLI 345


RESULT 15
AAW75995
ID   AAW75995 standard; Protein; 1261 AA.
XX
AC   AAW75995;
XX
DT   18-NOV-1998  (first entry)
XX
DE   GTPase activating protein (GAP), PARG.
XX
KW   PARG; GTPase activating protein; GAP; PTPL1 polypeptide; phosphatase;
KW   Rho family signal transduction; cancer cell; proliferation; GAP domain;
KW   Ras-like GTPase; Rho; Rac; Cdc42; PDZ domain; signal transduction;
KW   intracellular protein tyrosine phosphatase; mast cell secretion;
KW   pharmacological agent; modulator; diagnosis.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          193..509
FT                   /note= "ZPH domain"
FT   Domain          613..652
FT                   /note= "cysteine-rich domain"
FT   Domain          658..898
FT                   /note= "GAP domain"
XX
PN   WO9837196-A1.
XX
PD   27-AUG-1998.
XX
PF   19-FEB-1998;   98WO-US03323.
XX
PR   25-FEB-1997;   97US-0805583.
XX
```

PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Aspenstrom P, Franzen P, Gonez LJ, Heldin C, Hellman U;
PI Saras J;
XX
DR WPI; 1998-467559/40.
DR N-PSDB; AAV55846.
XX
PT Polypeptide, PARG, with GTPase activating activity - useful e.g.
PT therapeutically to reduce Rho family signal transduction so reduce
PT cancer cell proliferation, modulate mast cell secretion
XX
PS Claim 18; Pages 41-48; 93pp; English.
XX
CC This represents a human GTPase activating protein (GAP), PARG. The
CC specification states that the unique PARG nucleic acid molecule
CC specifically excludes known sequences shown in AAV55847 to AAV55850.
CC Expression vectors comprising the PARG nucleic acid can be used to
CC transform host cells for the recombinant production of the protein. The
CC PARG polypeptides or complexes comprising PTPL1 polypeptide having
CC phosphatase activity (AAW75999) and PARG polypeptide can be administered
CC to mammalian cells to reduce Rho family signal transduction. The PARG
CC polypeptides can be administered to cancer cells to reduce proliferation
CC (which is increased by Rho family signal transduction). Alternatively,
CC dominant-negative variants of the polypeptide (e.g. with a GAP domain
CC inactivated) can be used to increase Rho family signal transduction.
CC PARG was determined to have GAP activity in vitro, which was strongest on
CC Rho but also detected on Ras-like GTPases, Rac and Cdc42, and bound
CC specifically to the fourth PDZ domain (PDZ4) of known intracellular
CC protein tyrosine phosphatase PTPL1, involved in signal transduction.
CC PARG polypeptides and polynucleotides are useful to produce modulators of
CC PARG GTPase activity useful to reduce Rho family signal transduction or
CC modulate mast cell secretion. The polypeptides are also useful for
CC identifying pharmacological agents useful for disease diagnosis and
CC treatment.
XX
SQ Sequence 1261 AA;

Query Match 9.3%; Score 301; DB 19; Length 1261;
Best Local Similarity 22.4%; Pred. No. 1.1e-15;
Matches 153; Conservative 120; Mismatches 249; Indels 160; Gaps 33;

```
Qy          9 RNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHE-LGKYKDLLMKAETERSA 67
              | | |: :::||   | :|:  :    : |: |   : |  | | : : |: : :    |
Db        372 RRLEEEALQKVE---EADELYKV-CVTNVEERRNDVENTKREILAQLRTLVFQCDLTLKA 427

Qy         68 LDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEEQK---SA 124
              : | | ::     :  | ::      ||    |   | :    |     :||:|   :
Db        428 VTVNLFHMQHLQAASLADRLQSLCGSAKLYDPGQEYSEFVKATNS-----TEEEKVDGNV 482

Qy        125 LAFLNRGQPSSSNAGNKRLSTI---DESGSILSDISFDKTDESLDWDSSLVKTFKLKKRE 181
                ||  |||     |     :   | |  |  |   :  | :     | ::::
Db        483 NKHLNSSQPSGFGPANSLEDVVRLPDSSNKIEEDRCSNSADIT---GPSFIRSWTFGMFS 539

Qy        182 KRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVPY 241
              ||          |   ::||: |      |||         | |
Db        540 DSEST---------GGSSESRSLDS------ESI-------SPGD--------------- 562

Qy        242 WTRSRRKTGTLQPWNSDSTLNSRQLEPRTET--DSVGTPQS---NGGMRLHDFVSKTVIK 296
              : |   :| :    :|   |: |:    :||  :|:|| :    :     | |  : :
Db        563 FHRKLPRTPSSGTMSSADDLDEREPPSPSETGPNSLGTFKKTLMSKAALTHKF--RKLRS 620

Qy        297 PESCVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPC----IP---TLIGTPVKIGEG 349
              |  |  |   : |    ::| :| :| | :| :   : |    :|    | |
Db        621 PTKCRDCEGIVVF--QGVECEECLLVCHRKCLENLVIICGHQKLPGKIHLFG-------- 670

Qy        350 MLADF--VSQTSP-MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTV 406
              |:|  |::  |  || |:  | :||| | |   |:||: |             ::||
Db        671 --AEFTLVAKKEPDGIPFILKICASEIENRALCLQGIYRVCG-----------NKIKTE 716

Qy        407 PLLSKVD------DI-----HAICSLLKDFLRNLKEPLLTFRLNRAFMEAAE----ITDE 451
              |   ::      ||     | || :|| :|| | || : ||| : |:: |:    : :|
Db        717 KLCLALENGMHLVDISEFSSHDICDVLKLYLRQLPEPFILFRLYKEFIDLAKEIQHVNEE 776

Qy        452 DNS--------------------IAAMYQAVGELPQANRDTLAFLMIHLQRVA-QSPHTK 490
                :                    :      : :|| :| ::| ||::||:||   :   |
Db        777 QETKKNSLEDKKWPNMCIEINRILLKSKDLLRQLPASNFNSLHFLIVHLKRVVDHAEENK 836

Qy        491 MDVANLAKVFGPTIVAHAVPNPD--PVTMSQ--DIKRQPKVVERLLSLPLEYWSQFM--- 543
              |:  ||  :|||:::    | |   |:|:|   :   | ::|| |::     :|| :
Db        837 MNSKNLGVIFGPSLIR---PRPQTAPITISSLAEYSNQARLVEFLIT-----YSQKIFDG 888

Qy        544 -MVEQENIDPLHVIENS---NAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVR 599
               :  |: :  : |::         :|:  ||: |:     : :  :  : | | : :|
Db        889 SLQPQDVMCSIGVVDQGCFPKPLLSPEERDIERSMKSLFFSSKEDIHTSESESKIFERAT 948

Qy        600 STLTKNTPRFGSKSKSATNLGR 621
              |          |    :   ||:
Db        949 S--------FEESERKQNALGK 962
```

Search completed: July 3, 2003, 09:54:08
Job time : 76 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 3, 2003, 09:52:57 ; Search time 43 Seconds
(without alignments)
1412.952 Million cell updates/sec

Title: US-09-881-736-2
Perfect score: 3243
Sequence: 1 MDTMMLNVRNLFEQLVRRVE.....SKSATNLGRQGNFFASPMLK 632

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3243	100.0	632	2	D59430	Rac GTPase activat
2	547.5	16.9	681	2	T23454	hypothetical prote
3	477	14.7	383	2	B48122	GTPase-activating
4	477	14.7	384	2	A48122	GTPase-activating
5	348	10.7	1980	2	S54307	myosin heavy chain
6	334	10.3	2022	2	A59256	myosin-IXb [simila
7	329	10.1	334	2	S29128	N-chimerin - rat
8	322	9.9	299	2	S08242	N-chimerin - human
9	320	9.9	299	2	A43953	N-chimerin - commo
10	308	9.5	695	2	T39954	probable gtpase ac
11	301	9.3	1261	2	E59430	PTPL1-associated R
12	294	9.1	295	2	S29956	beta-chimerin - ra
13	294	9.1	443	2	B53764	beta2-chimerin, ce
14	293.5	9.1	466	2	A53764	beta2-chimerin, ce
15	291	9.0	970	2	D59435	Gem-interacting pr
16	290.5	9.0	814	2	F59430	GTPase regulator a
17	286.5	8.8	1944	2	A59438	KIAA1424 protein [
18	285.5	8.8	1165	2	D59433	C. elegans protein
19	282	8.7	859	2	A49307	98K GTPase-activat
20	281	8.7	813	2	B47485	ABR protein 2 - hu
21	281	8.7	822	2	A47485	ABR protein 1 - hu
22	280.5	8.6	733	2	S44876	ZC21.4 protein - C
23	275.5	8.5	1271	1	TVHUBR	bcr (breakpoint cl
24	270.5	8.3	1397	2	T46354	hypothetical prote
25	270.5	8.3	2548	2	E59435	myosin IXA [import
26	269.5	8.3	735	2	A59434	KIAA1501 protein [
27	267.5	8.2	2626	2	T31099	myosin-RhoGAP prot
28	260.5	8.0	802	2	H59434	oligophrenin 1, Rh
29	258	8.0	903	2	T00705	N-chimerin homolog
30	253.5	7.8	1846	2	T33079	hypothetical prote
31	247.5	7.6	655	2	A59430	hypothetical prote
32	247	7.6	863	2	T27958	hypothetical prote
33	245	7.6	1445	2	A59437	KIAA1204 protein [
34	240	7.4	969	2	T38478	RhoGAP/LIM domain
35	238	7.3	974	2	E59434	Rho GTPase activat
36	232	7.2	666	2	S29349	hypothetical prote
37	231.5	7.1	647	2	A57467	RalBP1 - rat
38	231	7.1	818	2	A59433	KIAA0672 protein [
39	229.5	7.1	574	2	T29005	hypothetical prote
40	229.5	7.1	655	2	F59435	ralA-binding prote
41	229	7.1	634	2	T27959	hypothetical prote
42	228.5	7.0	837	2	T19825	hypothetical prote
43	222.5	6.9	892	2	T40040	GTPase-activator p
44	221.5	6.8	316	2	T46471	hypothetical prote
45	217	6.7	512	2	E59437	F02569_2 protein [

ALIGNMENTS

RESULT 1
D59430
Rac GTPase activating protein 1 [imported] - huamn
C;Species: Homo sapiens (man)
C;Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 19-Jul-2002
C;Accession: D59430
R;Kawashima, T.; Kitamura, T.; Nosaka, T.; Hirose, K.
submitted to GenBank, December 1999
A;Description: Homo sapiens Rac GTPase activating protein 1 (RACGAP1), mRNA.
A;Reference number: D59430
A;Accession: D59430
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-632 <KAW>
A;Cross-references: GB:NP_037409; PID:g7019433; PIDN:NP_037409.1

Query Match 100.0%; Score 3243; DB 2; Length 632;
Best Local Similarity 100.0%; Pred. No. 1.9e-198;
Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60

Qy         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120

Qy        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180

Qy        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240

Qy        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300

Qy        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360

Qy        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420

Qy        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480
```

```
Qy        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540

Qy        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600

Qy        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              ||||||||||||||||||||||||||||||||
Db        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632


RESULT 2
T23454
hypothetical protein K08E3.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T23454
R;McMurray, A.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19743
A;Accession: T23454
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-681 <WIL>
A;Cross-references: EMBL:Z81568; PIDN:CAB04593.1; GSPDB:GN00021; CESP:K08E3.6
A;Experimental source: clone K08E3
C;Genetics:
A;Gene: CESP:K08E3.6
A;Map position: 3
A;Introns: 36/1; 73/2; 237/3; 361/3; 612/3

  Query Match             16.9%;  Score 547.5;  DB 2;  Length 681;
  Best Local Similarity   26.0%;  Pred. No. 4.8e-27;
  Matches  181;  Conservative  109;  Mismatches  284;  Indels  121;  Gaps   22;

Qy          7 NVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKW-------QRTDHELGKYKDLLM 59
              | |::|  ::       :  ::    |  : |  || |       :| : :: : :: |
Db         14 NSRHIFNMILNSQRPQFDIKDIGMFHLIDEIERLRKLWKDSEESKKRLNADMREAEEALA 73

Qy         60 KAETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSE 119
              ||  : :  |: :|  :  :   ::  :  : |    | : : ::: :      |  |::
Db         74 KARKKLAMFDIDVKDTQKHLRALMEENKALKLDLNVYETREKQLKDAMKNGIFNS--LTK 131

Qy        120 EQKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWD---SSLVKTFK 176
              | :   ||:        :|||:      :: |   |: |  :|:|    |  :
Db        132 EDRDQFKFLHE---PLVRTYSKRVQ--QRHPHLMEDTQDDEDDSEVDYDETGDSFEEVIH 186

Qy        177 LKK-REKRRSTS------------------------------RQFVDGPPG----PVKK 200
              |:  || |||::                            :|  |     | |:
Db        187 LRNGREVRRSSAAGNAVGGKRRSASAHAITAAANSKRSRSRVMTATIDEEPNEGGTPPKR 246

Qy        201 TRSIGSAVDQGNESIVAKTTVT------VPNDGGP-------------IEAVSTIETVP 240
              |  ||   |   :   || |        |   |                :: : :  |
Db        247 CRDDGSTPHQEMTTTTTTTTTTTIHNSRAQNQDPPRVSLHRQLTRRSLSCGSIPSCDQTP 306

Qy        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300
              | :   |      : |||: | |:  |   : || : :  || | |:   :     |
Db        307 GQTTNNIGLGMSSAILTKSTLDIRTLKRGTPAWTNGTTR-DIAMRPHTFIEAGIKAMRKC 365

Qy        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIP----------TLIGTPVKIGEGM 350
              |   :|    |:|||||  | |  | ::  |||||           | |     ||
Db        366 DKCATALKLA-TSMKCRDCHQVVHRSCCNKLHLPCIPRPKTMMTPKSALRGAKPGAGEFR 424

Qy        351 LADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLS 410
              | || :   ||||: |:|||  :| ||||: |:||: |  |||  | :: || |||| :
Db        425 LQDFCTSAKPMIPAAVIHCVVALEARGLTQEGIYRVPGQVRTVNVLLDE-LRSKTVPNVG 483

Qy        411 KVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEI--TDEDNSIAAMYQAVGELPQA 468
              : |:  |   || |||:||:||:     :  : || :  || ||   |: : : |||||
Db        484 -LHDVEVITDTLKRFLRDLKDPLIPRTSRQELIVAANLYSTDPDNGRLALNRVICELPQA 542
```

```
Qy        469 NRDTLAFLMIHLQRV-AQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQ-------D 520
              ||||||:| || ::| |||   ||:   :|::  | :: |      ||  ||.      |
Db        543 NRDTLAYLFIHWRKVIAQSSRNKMNCEAMARMVAPAVMGH------PVKQSQSQAIAGRD 596

Qy        521 IKRQPKVVERLLSLPLEYWSQFMMVEQENIDPLHVIENSNAFSTPQTPD----IKVSLLG 576
                : :  |      || :|:           |   ||   | :  |       |:||
Db        597 ATDCHRAMTALFEFDDVYWQRFLGTSA-------VSMASNQIETARHQDNFALCDRSILG 649

Qy        577 PVTTPEHQLLKTPSSSSLSQRVRSTLTKNTPRFGS 611
              ||||       :|::  |::   :|  :      ||
Db        650 PVTT-------SPATPLLARSANATRARGAHLLGS 677


RESULT 3
B48122
GTPase-activating protein Rac homolog, splice form clone pc1.7d - fruit fly (Drosophi
C;Species: Drosophila melanogaster
C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
C;Accession: B48122
R;Agnel, M.; Roder, L.; Vola, C.; Griffin-Shea, R.
Mol. Cell. Biol. 12, 5111-5122, 1992
A;Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal
A;Reference number: A48122; MUID:93024458; PMID:1406685
A;Accession: B48122
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-383 <AGN>
A;Note: sequence extracted from NCBI backbone (NCBIN:115662, NCBIP:115663)
C;Genetics:
A;Gene: FlyBase:rn
A;Cross-references: FlyBase:FBgn0003263

  Query Match             14.7%;  Score 477;  DB 2;  Length 383;
  Best Local Similarity   40.7%;  Pred. No. 6.5e-23;
  Matches  110;  Conservative   44;  Mismatches  102;  Indels   14;  Gaps    6;

Qy        279 QSNGG-MRLHDFVSKT-VIKPESCVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCI 336
              ||: | :| |:|  |:      :|| | |||:|   ||:|| | :  |  |  :  : ||
Db         78 QSHSGLLREHNFKIKSYYYNVGNCVHCRKRIRFAMASLRCRACPLRCHIGCCRQLTVNCI 137

Qy        337 P-TLIGTPVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKE 395
              |   |||       | |:|:  : :||:|:::|||| ||| ||| : ||||:|      |
Db        138 PQPQIGTK----RGCLSDYAPRVAPMVPALIVHCVTEIEARGLQQEGLYRVSSTREKCKR 193

Qy        396 LKEKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSI 455
              |: | || |: | |     | | :|  :|||| |  ||:      | | ||     |
Db        194 LRRKLLRGKSTPHLGN-KDTHTLCCCVKDFLRQLVHPLIPIYHRRDFEEATRHEDRLAVE 252

Qy        456 AAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPV 515
               |:| || || ||:|||||:||:| |::|:||  :| | ||| :| ||:        |
Db        253 MAVYLAVLELHQAHRDTLAYLMLHWQKIAESPAVRMTVNNLAVIFAPTLFG------DLD 306

Qy        516 TMSQDIKRQPKVVERLLSLPLEYWSQFMMV 545
                 :::    :|:: || :|  :||||: |
Db        307 LTLENVVTWQRVLKVLLLMPAGFWSQFLEV 336


RESULT 4
A48122
GTPase-activating protein Rac homolog, splice form clone pc1.7 - fruit fly (Drosophil
C;Species: Drosophila melanogaster
C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
C;Accession: A48122
R;Agnel, M.; Roder, L.; Vola, C.; Griffin-Shea, R.
Mol. Cell. Biol. 12, 5111-5122, 1992
A;Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal
A;Reference number: A48122; MUID:93024458; PMID:1406685
A;Accession: A48122
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-384 <AGN>
A;Note: sequence extracted from NCBI backbone (NCBIN:115660, NCBIP:115661)
C;Genetics:
```

```
A;Gene: FlyBase:rn
A;Cross-references: FlyBase:FBgn0003263

  Query Match           14.7%;  Score 477;  DB 2;  Length 384;
  Best Local Similarity  40.7%;  Pred. No. 6.5e-23;
  Matches  110;  Conservative   44;  Mismatches  102;  Indels   14;  Gaps    6;

Qy        279 QSNGG-MRLHDFVSKT-VIKPESCVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCI 336
              ||: | :| |:|  |:      :|| | |||:|   ||:|| | :  |  |  :  : ||
Db         78 QSHSGLLREHNFKIKSYYYNVGNCVHCRKRIRFAMASLRCRACPLRCHIGCCRQLTVNCI 137

Qy        337 P-TLIGTPVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKE 395
              |   |||      | |:|:  : :||:|:::|||| ||| ||| : ||||:|      |
Db        138 PQPQIGTK----RGCLSDYAPRVAPMVPALIVHCVTEIEARGLQQEGLYRVSSTREKCKR 193

Qy        396 LKEKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSI 455
              |: | || |: | |    | | :|  :|||| |  ||:   .| | ||    |
Db        194 LRRKLLRGKSTPHLGN-KDTHTLCCCVKDFLRQLVHPLIPIYHRRDFEEATRHEDRLAVE 252

Qy        456 AAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPV 515
               |:| || || ||:|||||:||:| |::|:||  :| | ||| :| ||:        |
Db        253 MAVYLAVLELHQAHRDTLAYLMLHWQKIAESPAVRMTVNNLAVIFAPTLFG------DLD 306

Qy        516 TMSQDIKRQPKVVERLLSLPLEYWSQFMMV 545
              :::    :|:: || :|  :||||: |
Db        307 LTLENVVTWQRVLKVLLLMPAGFWSQFLEV 336


RESULT 5
S54307
myosin heavy chain - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 02-Mar-2001
C;Accession: S54307
R;Reinhard, J.; Scheel, A.A.; Diekmann, D.; Hall, A.; Ruppert, C.; Baehler, M.
EMBO J. 14, 697-704, 1995
A;Title: A novel type of myosin implicated in signalling by rho family GTPases.
A;Reference number: S54307; MUID:95188874; PMID:7882973
A;Accession: S54307
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1980 <REI>
A;Cross-references: EMBL:X77609; NID:g639998; PIDN:CAA54700.1; PID:g639999
C;Superfamily: rat myosin heavy chain; myosin motor domain homology; protein kinase C zi
C;Keywords: nucleotide binding; P-loop
F;149-942/Domain: myosin motor domain homology <MMOT>
F;239-246/Region: nucleotide-binding motif A (P-loop)
F;1593-1641/Domain: protein kinase C zinc-binding repeat homology <KZN>

  Query Match           10.7%;  Score 348;  DB 2;  Length 1980;
  Best Local Similarity  22.1%;  Pred. No. 1e-13;
  Matches  160;  Conservative  105;  Mismatches  262;  Indels  196;  Gaps   25;

Qy         82 EIKRRQRAEADCEKLERQIQLIR--EMLMCDTSGSIQLSEEQKSALAFLNRGQP---SSS 136
              :|:| |    | |:|   ::: |  ::     | |: |||: ::| | |   :    | |
Db       1272 QIQRYQ--HPDTERLATAVEIWRGKKLASAMLSQSLDLSEKPRTAGAALTPTEERRISFS 1329

Qy        137 NAGNKRLSTIDESGSILSDISFDK-------------------TDESLDWDSSLVKTFK 176
              :   :|| :  |  :  |:|  |                   |    |  |:  : |
Db       1330 TSDVSKLSPVKTSTEVDGDLSAKKPAGHKKKSEDPSAGPDAGLPTGSQGDSKSAFKRLFL 1389

Qy        177 LKKREKRRS-----------------TSRQFVDGP--------------PGPVKKTRSI 204
              | ::|: |                   :|: :| |              |   : | :
Db       1390 HKAKDKKPSLEGVEETEGSGGQAAQEAPARKTLDVPSSQQHRHTTGEKPLKGKKNRNRKV 1449

Qy        205 G--SAVDQGNESIVAKTT-----------VTVPND----GGPIEA------------VS 234
              |  :  ::  ||:  | |            :   ||      |||:            :
Db       1450 GQITVSEKWRESVFRKITNANELKFLDEFLLNKVNDLRSQKTPIESLFIEATERFRSNIK 1509

Qy        235 TIETVP-------------------------------------------YWTRSRRKTGT 251
              |: :||                                          :|||  ||
Db       1510 TMYSVPNGKIHVGYKDLMENYQIVVSNLAAERGEKDTNLVLNVFQSLLDEFTRSYNKT-D 1568

Qy        252 LQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGK 311
              :|    :    |: |   :       : ||    | | |  |  |:||  |   |
Db       1569 FEPVKGKAQKKKRKQERAVQ-------EHNG----HVFASYQVNIPQSCEQCLSYIWLMD 1617

Qy        312 LSLKCRDCRVVSHPECRDRCPLPCIPT------LIGTPVKIGEGMLADFVSQTSPMIPSI 365
              :| |  |:: | :|  :   |  |       |   |      |:  | ::      :| :
Db       1618 KALLCSVCKMTCHKKCVHKIQSYCSYTGRRKSELGAEPGHF--GVCVDSLTSDKASVPIV 1675

Qy        366 VVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDD--IHAICSLLK 423
              :   :  :|  ||   |||| ||     :||::    ::|  |   |::|  |||| :||
Db       1676 LEKLLEHVEMHGLYTEGLYRKSGAANRTRELRQ---ALQTDPATVKLEDFPIHAITGVLK 1732

Qy        424 DFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRV 483
              :|| | |||:||     |: | |: ::   :||:|  :  ||:|| :|  |: || :|
Db       1733 QWLRELPEPLMTFAQYGDFLRAVELPEKQEQLAAIYAVLDHLPEANHTSLERLIFHLVKV 1792

Qy        484 A-QSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWSQF 542
              |      :|    || :| | ::     | ||:|  :|: :    || |:   :  :
Db       1793 ALLEDVNRMSPGALAIIFAPCLL-RCPDNSDPLTSMKDVLKITTCVEMLIKEQMRKYK-- 1849

Qy        543 MMVEQENIDPLHVIEN----------SNA-------FSTPQTPDIKVSLLGPVTTPEHQL 585
              |: | |: |   |:            ||       ||:|       |   ||     :|
Db       1850 --VKMEEINHLEAAESIAFRRLSLLRQNAPWPLKLGFSSPYEGVRTKSPRTPVVQDLEEL 1907

Qy        586 LKTPSSSS----------LSQRVRS------TLTKNTPRFGSKSKSATNLGRQGNFFASP 629
              |  ::           | :|::|       :|   |    :     ||  : :
Db       1908 GALPEEAAGGDEDREKEILMERIQSIKEEKEDITYRLPELDPRGSDEENLDSETSASTES 1967

Qy        630 MLK 632
              :|:
Db       1968 LLE 1970


RESULT 6
A59256
myosin-IXb [similarity] - human
C;Species: Homo sapiens (man)
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 02-Mar-2001
C;Accession: A59256; I61700
R;Wirth, J.A.; Jensen, K.A.; Post, P.L.; Bement, W.M.; Mooseker, M.S.
J. Cell Sci. 109, 653-661, 1996
A;Title: Human myosin-IXb, an unconventional myosin with a chimerin-like rho/rac GTPa
A;Reference number: A59256; MUID:97063843; PMID:8907710
A;Accession: A59256
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-2022 <WIR>
A;Cross-references: GB:U42391; NID:g1147782; PIDN:AAC50402.1; PID:g1147783
R;Bement, W.M.; Hasson, T.; Wirth, J.A.; Cheney, R.E.; Mooseker, M.S.
Proc. Natl. Acad. Sci. U.S.A. 91, 6549-6553, 1994
A;Title: Identification and overlapping expression of multiple unconventional myosin
A;Reference number: A55758; MUID:94294418; PMID:8022818
A;Accession: I61700
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 234-322 <RES>
A;Cross-references: GB:L29149; NID:g457257; PIDN:AAA20912.1; PID:g531142
C;Genetics:
A;Gene: GDB:MYO9B; OMIM:602129
A;Map position: 19p13.1
C;Superfamily: rat myosin heavy chain; myosin motor domain homology; protein kinase C
C;Keywords: nucleotide binding; P-loop
F;149-941/Domain: myosin motor domain homology #status atypical <MMO>
F;239-246/Region: nucleotide-binding motif A (P-loop)

  Query Match           10.3%;  Score 334;  DB 2;  Length 2022;
  Best Local Similarity  23.0%;  Pred. No. 8.3e-13;
  Matches  147;  Conservative  111;  Mismatches  266;  Indels  116;  Gaps   26;

Qy         22 LSEGNEVQ--------FIQLAKDFEDFRKKWQ---RTDHELGKYKDLLMKAETERSALDV 70
              || |::|         |:   ||     ||:      : |    :::: | | :  |:
Db       1404 LSPGSQVDSKSTFKRLFLHKTKD-----KKYSLEGAEELENAVSGHVVLEATTMKKGLEA 1458
```

```
Qy        71 ----KLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEEQKSALA 126
Db      1459 PSGQQHRHAAGE------KRTKEPGGKGKKNRNVKI----------GKITVSEKWRESVF 1502

Qy       127 FLNRGQPSSSNAGNKRLSTIDE-SGSILSDISFDKTDESLDWDSSLVKTFKLKKREKRRS 185
Db      1503 ----RQITNAN----ELKYLDEFLLNKINDLRSQKTP---------IESLFIEATEKFRS 1545

Qy       186 TSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIET-VPYWTR 244
Db      1546 NIKTMYSVPNGKI----HVGYKDLMENYQIVVSNLATERGQKDTNLVLNLFQSLLDEFTR 1601

Qy       245 SRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCG 304
Db      1602 GYTK-------NDFEPVKQSKAQKKKRKQERAVQEHNG----HVFASYQVSIPQSCEQCL 1650

Qy       305 KRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTL--IGTP-VKIGE-GMLADFVSQTSP 360
Db      1651 SYIWLMDKALLCSVCKMTCHKKCVHKIQSHCSYTYGRKGEPGAEPGHFGVCVDSLTSDKA 1710

Qy       361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDD--IHAI 418
Db      1711 SVPIVLEKLLEHVEMHGLYTEGLYRKSGAANRTRELRQ---ALQTDPAAVKLENFPIHAI 1767

Qy       419 CSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMI 478
Db      1768 TGVLKQWLRELPEPLMTFAQYGDFLRAVELPEKQEQLAAIYAVLEHLPEANHNSLERLIF 1827

Qy       479 HLQRVA-QSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLE 537
Db      1828 HLVKVALLEDVNRMSPGALAIIFAPCLL-RCPDNSDPLTSMKDVLKITTCVEMLIKEQMR 1886

Qy       538 YWSQFMMVEQENIDPLHVIEN----------SNAFSTPQTPD--------IKVSLLGPVT 579
Db      1887 KYK----VKMEEISQLEAAESIAFRRLSLLRQNANKSPKTREPAGGAGRLLTTSRVSP-- 1940

Qy       580 TPEHQLLKTPSSSSLSQRVRST----------LTKNTPR 608
Db      1941 SPSTRNLALGSWRSAALRTRGTGRPARPGRARALRRRPPR 1980
```

RESULT 7
S29128
N-chimerin - rat
N;Alternate names: GTPase-activating protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 25-Feb-1994 #sequence_revision 19-Oct-1995 #text_change 08-Oct-1999
C;Accession: S29128; S25152
R;Lim, H.H.; Michael, G.J.; Smith, P.; Lim, L.; Hall, C.
Biochem. J. 287, 415-422, 1992
A;Title: Developmental regulation and neuronal expression of the mRNA of rat n-chimaerin
A;Reference number: S29128; MUID:93074974; PMID:1445199
A;Accession: S29128
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-334 <LIM>
A;Cross-references: EMBL:X67250; NID:g55939; PIDN:CAA47672.1; PID:g55940
R;Lim, H.; Michael, G.J.; Smith, P.; Lim, L.; Hall, C.
submitted to the EMBL Data Library, July 1992
A;Description: Rat n-chimaerin a p21rac GAP:cDNA sequence developmental regulation and n
A;Reference number: S25152
A;Accession: S25152
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-334 <LIM2>
A;Cross-references: EMBL:X67250; NID:g55939; PIDN:CAA47672.1; PID:g55940
C;Superfamily: protein kinase C zinc-binding repeat homology
F;81-130/Domain: protein kinase C zinc-binding repeat homology <KZ2>

  Query Match 10.1%; Score 329; DB 2; Length 334;
  Best Local Similarity 32.4%; Pred. No. 1.4e-13;
  Matches 91; Conservative 46; Mismatches 118; Indels 26; Gaps 8;

```
Qy       259 STLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLS--LKC 316
Db        66 ATLKENEQIPKYE------------KVHNFKVHTFRGPHWCEYCA-NFMWGLIAQGVKC 111

Qy       317 RDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQR 376
Db       112 ADCGLNVHKQCSKMVPNDCKPDLKHVK-KVYSCDLTTLVKAHITKRPMVVDMCIREIESR 170

Qy       377 GLTETGLYRISGCDRTVKELKEKFLR--VKTVPLLSKVDDIHAICSLLKDFLRNLKEPLL 434
Db       171 GLNSEGLYRVSGFSDLIEDVKMAFDRDGEKADISVNMYEDINIITGALKLYFRDLPIPLI 230

Qy       435 TFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTK---M 491
Db       231 TYDAYPKFIESAKIVDPDEQLETLHEALRSLPPAHCETLRYLMAHLKRV--TLHEKENLM 288

Qy       492 DVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLL 532
Db       289 SAENLGIVFGPTLMRS--PELDPMAALNDIRYQRLVVELLI 327
```

RESULT 8
S08242
N-chimerin - human
N;Alternate names: GTPase-activating protein
C;Species: Homo sapiens (man)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 05-Nov-1999
C;Accession: S08242
R;Hall, C.; Monfries, C.; Smith, P.; Lim, H.H.; Kozma, R.; Ahmed, S.; Vanniasingham,
J. Mol. Biol. 211, 11-16, 1990
A;Title: Novel human brain cDNA encoding a 34,000 M(r) protein n-chimaerin, related t
A;Reference number: S08242; MUID:90133942; PMID:2299665
A;Accession: S08242
A;Molecule type: mRNA
A;Residues: 1-299 <HAL>
A;Cross-references: EMBL:X51408; NID:g35012; PIDN:CAA35769.1; PID:g35013
C;Superfamily: protein kinase C zinc-binding repeat homology
F;46-95/Domain: protein kinase C zinc-binding repeat homology <KZ2>

  Query Match 9.9%; Score 322; DB 2; Length 299;
  Best Local Similarity 32.0%; Pred. No. 3.3e-13;
  Matches 90; Conservative 48; Mismatches 117; Indels 26; Gaps 8;

```
Qy       259 STLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLS--LKC 316
Db        31 ATLKENEQIPKYE------------KIHNFKVHTFRGPHWCEYCA-NFMWGLIAQGVKC 76

Qy       317 RDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQR 376
Db        77 ADCGLNVHKQCSKMVPNDCKPDLKHVK-KVYSCDLTTLVKAHTTKRPMVVDMCIREIESR 135

Qy       377 GLTETGLYRISGCDRTVKELKEKFLR--VKTVPLLSKVDDIHAICSLLKDFLRNLKEPLL 434
Db       136 GLNSEGLYRVSGFSDLIEDVKMAFDRDGEKADISVNMYEDINIITGALKLYFRDLPIPLI 195

Qy       435 TFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTK---M 491
Db       196 TYDAYPKFIESAKIMDPDEQLETLHEALKLLPPAHCETLRYLMAHLKRV--TLHEKENLM 253

Qy       492 DVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLL 532
Db       254 NAENLGIVFGPTLMRS--PELDAMAALNDIRYQRLVVELLI 292
```

RESULT 9
A43953
N-chimerin - common canary
N;Alternate names: protein kinase C homolog [misidentification]; song control circuit
C;Species: Serinus canaria (common canary)
C;Date: 26-Aug-1999 #sequence_revision 26-Aug-1999 #text_change 04-Feb-2000
C;Accession: A43953
R;George, J.M.; Clayton, D.F.
Brain Res. Mol. Brain Res. 12, 323-329, 1992

```
A;Title: Differential regulation in the avian song control circuit of an mRNA predicting
A;Reference number: A43953; MUID:92252621; PMID:1374499
A;Accession: A43953
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-299 <GEO>
A;Cross-references: GB:S98891; NID:g249117; PIDN:AAB22136.1; PID:g249118
A;Experimental source: HVC-associated telencephalon
A;Note: sequence extracted from NCBI backbone (NCBIN:98891, NCBIP:98892)
C;Superfamily: protein kinase C zinc-binding repeat homology
F;46-95/Domain: protein kinase C zinc-binding repeat homology <KZ2>

  Query Match             9.9%;  Score 320;  DB 2;  Length 299;
  Best Local Similarity   31.6%;  Pred. No. 4.4e-13;
  Matches   92;  Conservative   47;  Mismatches  106;  Indels   46;  Gaps   10;

Qy        259 STLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLS--LKC 316
              :||   :  |: |             ::|:|   |   |  |  |     :| ::  :||
Db         31 ATLKENEHVPKYE------------KVHNFKVHTFRGPHWCEYCA-NFMWGLIAQGVKC 76

Qy        317 RDCRVVSHPECRDRCPLPCIPTL----------IGTPVKIGEGMLADFVSQTSPMIPSIV 366
               || :  | :|    |  | | |          : | ||      | |  :     | :|
Db         77 ADCGLNVHKQCSKMVPNDCKPDLKHVKKVYSCDLTTLVK------AHFTKR-----PMVV 125

Qy        367 VHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLR--VKTVPLLSKVDDIHAICSLLKD 424
               |: ||| |||   ||||:||    ::::|  | |   |    ::  :||: |   ||
Db        126 DMCIREIESRGLNSEGLYRVSGFSDLIEDVKMAFDRDGEKADISVNMYEDINIITGALKL 185

Qy        425 FLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVA 484
              : |:|  ||:|:     |:|:|: || |  :  :::|:  || |: :|| :|| ||:||
Db        186 YFRDLPIPLITYDAYPKFIESAKTTDPDEQLEILHEALKLLPPAHCETLRYLMAHLKRV- 244

Qy        485 QSPHTK---MDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLL 532
               : | |   |   ||  |||||::    |  | :    ||: |  ||| |:
Db        245 -TLHEKENLMSAENLGIVFGPTLM--RAPELDAMAALNDIRYQRLVVEMLI 292


RESULT 10
T39954
probable gtpase activating protein homolog - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T39954
R;Xiang, Z.; Aves, S.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, January 1999
A;Reference number: Z21893
A;Accession: T39954
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-695 <XIA>
A;Cross-references: EMBL:AL035065; PIDN:CAA22624.1; GSPDB:GN00067; SPDB:SPBC23G7.08c
A;Experimental source: strain 972h-; cosmid c23G7
C;Genetics:
A;Gene: SPDB:SPBC23G7.08c
A;Map position: 2

  Query Match             9.5%;  Score 308;  DB 2;  Length 695;
  Best Local Similarity   23.7%;  Pred. No. 8.3e-12;
  Matches  137;  Conservative   87;  Mismatches  223;  Indels  132;  Gaps   23;

Qy         35 KDFEDFRKKWQRTDHE----LGKYKDLLMKAETE-RSALDVKLKHARNQVDVEIK---RR 86
              |  ::: || :    |    : | |     ||||   ||| |     |  | : |    :
Db        145 KSIKEYAKKQENAYLEAVMQMDKSKSRFKGAETEYNRALDNK-----NTGDSQKKVGFFK 199

Qy         87 QRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE-QKSALAF----------LNRGQPS 134
               :: |   ||| : :|  |    |    |: ::  ||  |           | |  |
Db        200 PKSNAQLTKLEDEARLKAENAESDMHSKIENAQNVQKQLLCIHRPNYIKQFFSLQREIES 259

Qy        135 SSNAGNKRLSTIDESGSILSDI---------------SFDKTDESLDWDSSLVKTFKLK 178
              |  |   | : : || ::|: :                : |  : : |:   ::    :|
Db        260 SLIANYLRYTKLCESNTLLNGLTIRPQKPTPTNCGLQHALDNINANTDFVQYVLHA-SIK 318

Qy        179 KREKRRST--SRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIE----A 232
               : :  |  |:  :  ||       | |      | :|  | |  :|:   |::    |
Db        319 HEDNKNPTDASKTKIIQPPSSYGTGSSAGKTNPPVNPTI--KVTAAIPS---PLQNTNPA 373

Qy        233 VSTIETVPYWTRSRRKTGTLQPWNSDSTLN--SRQLEPRTETDSVGTPQSNGGMRLHDFV 290
               ||       : :   : |  |  : :: :  :  |:| |  |: |:  |:
Db        374 PSTFPNPSVASPAFPNSSTSNPSTAPASASPLASTLKPSTANDTNGSSSSSSS------- 426

Qy        291 SKTVIKPESCVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCP---LPCIP----TLIGT- 342
                 | : |     :                 : |    : |   || :|    | | |
Db        427 -----NPRTSSPLASNAE-------------NKPPVAQQSPPVLLPTLPPIQTTTIQTS 467

Qy        343 ------PVKIGEGMLADFVSQTS-------------------------PMIPSIVVHCV 370
                   |  |      |     ||                          ||:||: |
Db        468 REVAPPPSSINSNRAASPFRPTSVSPQPSSPTKSLLFGARLDAIILREHSNIPNIVMQCT 527

Qy        371 NEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLL------SKVDDIHAICSLLKD 424
              :::|  ||   |:||:      |  |: :|   :  |||         :|:||:  |||
Db        528 SQVENFGLNLQGIYRVPSSSARVNMLRSQF---ENNPLLQLHTPEDYENDVHAVADLLKI 584

Qy        425 FLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVA 484
              | | |:|||:     | |::|  : ||     |:::|: :|| ||  |:  | ||| ::
Db        585 FFRELREPLIPDNHQRDFIDAGNVEDESRRRDAVHRAINDLPDANYSTIRHLTIHLAKIK 644

Qy        485 Q-SPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIK 522
              : |   ||   ||| ::||||:  |   |:  : |: |:
Db        645 ENSDVNKMSTNNLAIIWGPTIIKQAT-IPEISSFSRTIE 682


RESULT 11
E59430
PTPL1-associated RhoGAP protein 1 [imported] - huamn
C;Species: Homo sapiens (man)
C;Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 19-Jul-2002.
C;Accession: E59430
R;Saras, J.; Franzen, P.; Aspenstrom, P.; Hellman, U.; Gonez, L.J.; Heldin, C.-H.
submitted to GenBank, December 1997
A;Description: Homo sapiens PTPL1-associated RhoGAP 1 (PARG1), mRNA.
A;Reference number: E59430
A;Accession: E59430
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1261 <SAR>
A;Cross-references: GB:NP_004806; PID:g4758882; PIDN:NP_004806.1

  Query Match             9.3%;  Score 301;  DB 2;  Length 1261;
  Best Local Similarity   22.4%;  Pred. No. 5.4e-11;
  Matches  153;  Conservative   120;  Mismatches  249;  Indels  160;  Gaps   33;

Qy          9 RNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHE-LGKYKDLLMKAETERSA 67
              | | |: :::||   | :|:  :    : |: |   : |  | | | : : |: : :   |
Db        372 RRLEEEALQKVE---EADELYKV-CVTNVEERRNDVENTKREILAQLRTLVFQCDLTLKA 427

Qy         68 LDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEEQK---SA 124
              : | | | ::     :  | ::     ||    |   | :    |     :||:|   :
Db        428 VTVNLFHMQHLQAASLADRLQSLCGSAKLYDPGQEYSEFVKATNS-----TEEEKVDGNV 482

Qy        125 LAFLNRGQPSSSNAGNKRLSTI---DESGSILSDISFDKTDESLDWDSSLVKTFKLKKRE 181
                 ||  |||     |     :   | |  |  |   : | :     | ::::
Db        483 NKHLNSSQPSGFGPANSLEDVVRLPDSSNKIEEDRCSNSADIT---GPSFIRSWTFGMFS 539

Qy        182 KRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVPY 241
                 ||          |   ::||: |      |||         | |
Db        540 DSEST---------GGSSESRSLDS------ESI-------SPGD--------------- 562

Qy        242 WTRSRRKTGTLQPWNSDSTLNSRQLEPRTET--DSVGTPQS---NGGMRLHDFVSKTVIK 296
              : |   :| :   :|   |: |:    :||  :|:|| :     :   | | : :
Db        563 FHRKLPRTPSSGTMSSADDLDEREPPSPSETGPNSLGTFKKTLMSKAALTHKF--RKLRS 620

Qy        297 PESCVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPC----IP---TLIGTPVKIGEG 349
              |  |  |   : |    ::| :| :| | :| :   :  |    :|    | |
Db        621 PTKCRDCEGIVVF--QGVECEECLLVCHRKCLENLVIICGHQKLPGKIHLFG-------- 670

Qy        350 MLADF--VSQTSP-MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTV 406
```

```
                |:|  |::  |  || |:  | :||| | |   |:||: |            ::||
Db        671 --AEFTLVAKKEPDGIPFILKICASEIENRALCLQGIYRVCG-----------NKIKTE 716

Qy        407 PLLSKVD------DI-----HAICSLLKDFLRNLKEPLLTFRLNRAFMEAAE----ITDE 451
                |   ::      ||     | || :|| :|| | || : ||| : |:: |:    : :|
Db        717 KLCLALENGMHLVDISEFSSHDICDVLKLYLRQLPEPFILFRLYKEFIDLAKEIQHVNEE 776

Qy        452 DNS--------------------IAAMYQAVGELPQANRDTLAFLMIHLQRVA-QSPHTK 490
                :                    :      : :|| :| ::| ||::||:||   :   |
Db        777 QETKKNSLEDKKWPNMCIEINRILLKSKDLLRQLPASNFNSLHFLIVHLKRVVDHAEENK 836

Qy        491 MDVANLAKVFGPTIVAHAVPNPD--PVTMSQ--DIKRQPKVVERLLSLPLEYWSQFM--- 543
                |:  ||  :|||:::    | |   |:|:|   :   | ::|| |::     :|| :
Db        837 MNSKNLGVIFGPSLIR---PRPQTAPITISSLAEYSNQARLVEFLIT-----YSQKIFDG 888

Qy        544 -MVEQENIDPLHVIENS---NAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVR 599
                 :  |: :  : |::         :|:  ||: |:     : :  :  : | | : :|
Db        889 SLQPQDVMCSIGVVDQGCFPKPLLSPEERDIERSMKSLFFSSKEDIHTSESESKIFERAT 948

Qy        600 STLTKNTPRFGSKSKSATNLGR 621
                |       |    :    ||:
Db        949 S--------FEESERKQNALGK 962


RESULT 12
S29956
beta-chimerin - rat
N;Alternate names: GTPase-activating protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C;Accession: A45485; S31398; S29956
R;Leung, T.; How, B.E.; Manser, E.; Lim, L.
J. Biol. Chem. 268, 3813-3816, 1993
A;Title: Germ cell beta-chimaerin, a new GTPase-activating protein for p21rac, is specif
A;Reference number: A45485; MUID:93179371; PMID:8440677
A;Accession: A45485
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-295 <LEU>
A;Cross-references: EMBL:X69489; NID:g57526; PIDN:CAA49244.1; PID:g57527
A;Experimental source: testis
A;Note: sequence extracted from NCBI backbone (NCBIP:125731)
R;Leung, T.; How, B.E.; Manser, E.; Lim, L.
submitted to the EMBL Data Library, November 1992
A;Description: Germ cell beta-chimaerin, a new GTPase-activating protein for p21rac, is
A;Reference number: S31398
A;Accession: S31398
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-295 <LE2>
A;Cross-references: EMBL:X69462; NID:g55816; PID:g55817
C;Superfamily: protein kinase C zinc-binding repeat homology
C;42-91/Domain: protein kinase C zinc-binding repeat homology <KZ2>

  Query Match             9.1%;  Score 294;  DB 2;  Length 295;
  Best Local Similarity   29.3%;  Pred. No. 1.9e-11;
  Matches   81;  Conservative   52;  Mismatches  111;  Indels   32;  Gaps    8;

Qy        285 RLHDFVSKTVIKPESCVPCGKRIKFGKLS--LKCRDCRVVSHPECRDRCPLPCIPTLIGT 342
                : |:|   |   |  | |     :| ::  ::| || :  | :|    |  | | |
Db         40 KTHNFKVHTFRGPHWCEYCA-NFMWGLIAQGVRCSDCGLNVHKQCSKHVPNDCQPDLKRI 98

Qy        343 PVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLR 402
                |:    |   |   :   | :|  |: ||| |||   ||||:||    ::::|  | |
Db         99 K-KVYCCDLTTLVKAHNTQRPMVVDICIREIEARGLKSEGLYRVSGFTEHIEDVKMAFDR 157

Qy        403 VKTVPLLSKVD-------DIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSI 455
                | |       ||: |   || : |:|  |::|:     |:|||:|:: |  :
Db        158 DG-----EKADISANIYPDINIITGALKLYFRDLPIPIITYDTYTKFIEAAKISNADERL 212

Qy        456 AAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTK-MDVANLAKVFGPTIVAHAVPNPDP 514
                |::: :  || |: :|| :|||||::|  :     |:  ||  |||||::    |
Db        213 EAVHEVLMLLPPAHYETLRYLMIHLKKVTMNEKDNLMNAENLGIVFGPTLM--RPPEDST 270

Qy        515 VTMSQDIKRQPKVVERLLSLPLEYWSQFMMVEQENI 550
                :|   |:: |  :|:             :::| |::
Db        271 LTTLHDMRYQKLIVQ------------ILIENEDV 293


RESULT 13
B53764
beta2-chimerin, cerebellar - rat (fragment)
N;Alternate names: GTPase-activating protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Mar-1999
C;Accession: B53764
R;Leung, T.; How, B.E.; Manser, E.; Lim, L.
J. Biol. Chem. 269, 12888-12892, 1994
A;Title: Cerebellar beta2-chimaerin, a GTPase-activating protein for p21 Ras-related
A;Reference number: A53764; MUID:94230370; PMID:8175705
A;Accession: B53764
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-443 <LEU>
C;Superfamily: protein kinase C zinc-binding repeat homology; SH2 homology
F;34-114/Domain: SH2 homology <SH2>
F;190-239/Domain: protein kinase C zinc-binding repeat homology <KZ2>

  Query Match             9.1%;  Score 294;  DB 2;  Length 443;
  Best Local Similarity   29.3%;  Pred. No. 3.4e-11;
  Matches   81;  Conservative   52;  Mismatches  111;  Indels   32;  Gaps    8;

Qy        285 RLHDFVSKTVIKPESCVPCGKRIKFGKLS--LKCRDCRVVSHPECRDRCPLPCIPTLIGT 342
                : |:|   |   |  | |     :| ::  ::| || :  | :|    |  | | |
Db        188 KTHNFKVHTFRGPHWCEYCA-NFMWGLIAQGVRCSDCGLNVHKQCSKHVPNDCQPDLKRI 246

Qy        343 PVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLR 402
                |:    |   |   :   | :|  |: ||| |||   ||||:||    ::::|  | |
Db        247 K-KVYCCDLTTLVKAHNTQRPMVVDICIREIEARGLKSEGLYRVSGFTEHIEDVKMAFDR 305

Qy        403 VKTVPLLSKVD-------DIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSI 455
                | |       ||: |   || : |:|  |::|:     |:|||:|:: |  :
Db        306 DG-----EKADISANIYPDINIITGALKLYFRDLPIPIITYDTYSKFIEAAKISNADERL 360

Qy        456 AAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTK-MDVANLAKVFGPTIVAHAVPNPDP 514
                |::: :  || |: :|| :|||||::|  :     |:  ||  |||||::    |
Db        361 EAVHEVLMLLPPAHYETLRYLMIHLKKVTMNEKDNLMNAENLGIVFGPTLM--RPPEDST 418

Qy        515 VTMSQDIKRQPKVVERLLSLPLEYWSQFMMVEQENI 550
                :|   |:: |  :|:             :::| |::
Db        419 LTTLHDMRYQKLIVQ------------ILIENEDV 441


RESULT 14
A53764
beta2-chimerin, cerebellar - human
N;Alternate names: GTPase-activating protein
C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000
C;Accession: A53764
R;Leung, T.; How, B.E.; Manser, E.; Lim, L.
J. Biol. Chem. 269, 12888-12892, 1994
A;Title: Cerebellar beta2-chimaerin, a GTPase-activating protein for p21 Ras-related
A;Reference number: A53764; MUID:94230370; PMID:8175705
A;Accession: A53764
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-466 <LEU>
A;Cross-references: GB:L29126; NID:g457229; PIDN:AAA19191.1; PID:g457230
C;Superfamily: protein kinase C zinc-binding repeat homology; SH2 homology
F;57-137/Domain: SH2 homology <SH2>
F;213-262/Domain: protein kinase C zinc-binding repeat homology <KZ2>

  Query Match             9.1%;  Score 293.5;  DB 2;  Length 466;
  Best Local Similarity   25.7%;  Pred. No. 4e-11;
  Matches   98;  Conservative   66;  Mismatches  155;  Indels   63;  Gaps   12;
```

```
Qy        199 KKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP------YWTRSRRKTGTL 252
Db        116 KRFESIHDLVTDGLITLYIETKAA--------EYISKMTTNPIYEHIGYATLLREKV--- 164

Qy        253 QPWNSDSTLNSRQLEPR---------TETDSVGTPQSNGGM----------RLHDFVSKT 293
Db        165 -----SRRLSRSKNEPRKTNVTHEEHTAVEKISSLVRRAALTHNDNHFNYEKTHNFKVHT 219

Qy        294 VIKPESCVPCGKRIKFGKLS--LKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGML 351
Db        220 FRGPHWCEYCA-NFMWGLIAQGVRCSDCGLNVHKQCSKHVPNDCQPDLKRIK-KVYCCDL 277

Qy        352 ADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLR--VKTVPLL 409
Db        278 TTLVKAHNTQRPMVVDICIREIEARGLKSEGLYRVSGFTEHIEDVKMAFDRDGEKADISA 337

Qy        410 SKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQAN 469
Db        338 NVYPDINIITGALKLYFRDLPIPVITYDTYSKFIDAAKISNADERLEAVHEVLMLLPPAH 397

Qy        470 RDTLAFLMIHLQRVAQSPHTK-MDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVV 528
Db        398 YETLRYLMIHLKKVTMNEKDNFMNAENLGIVFGPTLM--RPPEDSTLTTLHDMRYQKLIV 455

Qy        529 ERLLSLPLEYWSQFMMVEQENI 550
Db        456 Q------------ILIENEDV 464


RESULT 15
D59435
Gem-interacting protein [imported] - huamn
C;Species: Homo sapiens (man)
C;Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 19-Jul-2002
C;Accession: D59435
R;Aresta, S.; Beranger, F.; Berger, R.; de Gunzburg, J.
Submitted to GenBank, November 2000
A;Description: GMIP, a Gem interacting protein.
A;Reference number: D59435
A;Accession: D59435
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-970 <ARE>
A;Cross-references: GB:NP_057657; PID:g7706107; PIDN:NP_057657.1

  Query Match             9.0%;  Score 291;  DB 2;  Length 970;
  Best Local Similarity   22.6%;  Pred. No. 1.6e-10;
  Matches  188;  Conservative  112;  Mismatches  273;  Indels  260;  Gaps   40;

Qy         13 EQLVRRVEILSEGNEVQFIQLAKDFEDFRKK---W--QRTDHELGKYKDLLMKAETERSA 67
Db         81 EELDLRL-IRTKGGVDAALEYAKTWSRYAKELLAWTEKRASYELEFAKSTMKIAEAGKVS 139

Qy         68 LDVK------------LKH----------------------ARNQVDVE---------- 82
Db        140 IQQQSHMPLQYIYTLFLEHDLSLGTLAMETVAQQKRDYYQPLAAKRTEIEKWRKEFKEQW 199

Qy         83 IKRRQRAEADCEKLER-QIQLIR--EMLMCDTSGSIQLSEEQKSALAFLNRGQPSSSNAG 139
Db        200 MKEQKRMNEAVQALRRAQLQYVQRSEDLRARSQGSPEDSAPQAS---------PGPSKQQ 250

Qy        140 NKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKRE----KRRSTS--RQFV-D 192
Db        251 ERRRRSREEAQA--------KAQEAEALYQACVREANARQQDLEIAKQRIVSHVRKLVFQ 302

Qy        193 GPPGPVKKTRSI----GSAVDQGNESIVAKTTVTVPNDGG----------------PIE 231
Db        303 GDEVLRRVTLSLFGLRGAQAERGPRAFAALAECCAPFEPGQRYQEFVRALRPEAPPPPP 362

Qy        232 AVSTIETVPYWTRS-----RRKTGTLQ-----------PWNSDST----LNSRQLEPRTE 271
Db        363 AFSFQEFLPSLNSSPLDIRKKLSGPLPPRLDENSAEPGPWEDPGTGWRWQGTPGPTPGSD 422

Qy        272 TDSVG---------TPQSNGGMRLHDFVSKTVIKPESC----------VPCGKRIKF 309
Db        423 VDSVGGGSESRSLDSPTSSPGAGTRQLVKASSTGTESSDDFEERDPDLGDGLENGLGSPF 482

Qy        310 GKLSL----------------KCRDCR--VVSHPECRD---RCPLPCIPTLI-------- 340
Db        483 GKWTLSSAAQTHQLRRLRGPAKCRECEAFMVSGTECEECFLTCHKRCLETLLILCGHRRL 542

Qy        341 --GTPVKIGEGML---ADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKE 395
Db        543 PARTPL-FGVDFLQLPRDFPEE----VPFVVTKCTAEIEHRALDVQGIYRVSGSRVRVER 597

Qy        396 LKEKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDED--- 452
Db        598 LCQAFENGRALVELSG-NSPHDVSSVLKRFLQELTEPVIPFHLYDAFISLAKTLHADPGD 656

Qy        453 ---------NSIAAMYQAVGELPQANRDTLAFLMIHLQRV-AQSPHTKMDVANLAKVFGP 502
Db        657 DPGTPSPSPDVIRSLKTLLVQLPDSNYNTLRHLVAHLFRVAARFMENKMSANNLGIVFGP 716

Qy        503 TIVAHAVPNPD--------PVTMSQDIKRQPKVVERLLSLPLEYWSQFMMVEQENIDPLH 554
Db        717 TLLR----PPDGPRAASAIPVTCLLDSGHQAQLVE-----------FLIVHYEQIFGMD 760

Qy        555 VIENSNAFSTPQTPDIKVSLLGPVTT----------PEHQ---LLKTPS-----SSSLSQ 596
Db        761 ELPQ----ATEPPPQDFSPAPGPLTTSSQPPPPHLDPDSQPPVLASDPGPDPQHHSTLEQ 816

Qy        597 RVRSTLTK-NTPRFGSKSKSATN----------------LGRQ--GNFFASP 629
Db        817 HPTATPTEIPTPQSDQREDVAEDTKDGGGEVSSQGPEDSLLGTQSRGHFSRQP 869


Search completed: July  3, 2003, 09:57:41
Job time : 46 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 3, 2003, 09:44:41 ; Search time 25 Seconds
(without alignments)
1048.522 Million cell updates/sec

Title: US-09-881-736-2
Perfect score: 3243
Sequence: 1 MDTMMLNVRNLFEQLVRRVE.....SKSATNLGRQGNFFASPMLK 632

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	477	14.7	384	1	RN_DROME	P40809 drosophila
2	348	10.7	1980	1	MY9B_RAT	Q63358 rattus norv
3	339.5	10.5	2114	1	MY9B_MOUSE	Q9qy06 mus musculu
4	329	10.1	334	1	CHIN_RAT	P30337 rattus norv
5	328	10.1	334	1	CHIN_MOUSE	Q91v57 mus musculu
6	323.5	10.0	2158	1	MY9B_HUMAN	Q13459 homo sapien
7	322	9.9	459	1	CHIN_HUMAN	P15882 homo sapien
8	294	9.1	295	1	CHIO_RAT	Q03070 rattus norv
9	293.5	9.1	468	1	CHIO_HUMAN	P52757 homo sapien
10	290.5	9.0	814	1	OPHL_HUMAN	Q9una1 homo sapien
11	282	8.7	859	1	ABR_HUMAN	Q12979 homo sapien
12	280.5	8.6	747	1	YN54_CAEEL	P34588 caenorhabdi
13	275.5	8.5	1271	1	BCR_HUMAN	P11274 homo sapien
14	260.5	8.0	802	1	OPH1_HUMAN	O60890 homo sapien
15	251.5	7.8	986	1	RHG6_MOUSE	O54834 mus musculu
16	244	7.5	802	1	OPH1_MOUSE	Q99j31 mus musculu
17	238	7.3	974	1	RHG6_HUMAN	O43182 homo sapien
18	232	7.2	666	1	RGD1_YEAST	P38339 saccharomyc
19	228.5	7.0	837	1	YLE5_CAEEL	P46941 caenorhabdi
20	206.5	6.4	1275	1	YAU9_SCHPO	Q10164 schizosacch
21	202.5	6.2	946	1	RHG4_HUMAN	P98171 homo sapien
22	197.5	6.1	638	1	Y053_HUMAN	P42331 homo sapien
23	192	5.9	718	1	RHG8_HUMAN	Q9nsg0 homo sapien
24	191	5.9	425	1	RHG8_MOUSE	Q9cxp4 mus musculu
25	188.5	5.8	439	1	RHG1_HUMAN	Q07960 homo sapien
26	187	5.8	2167	1	BEM2_YEAST	P39960 saccharomyc
27	180	5.6	622	1	3BP1_HUMAN	Q9y313 homo sapien
28	171	5.3	650	1	Y411_HUMAN	O43295 homo sapien
29	169.5	5.2	1007	1	RGA1_YEAST	P39083 saccharomyc
30	169	5.2	601	1	3BP1_MOUSE	P55194 mus musculu
31	160.5	4.9	1017	1	LRG1_YEAST	P35688 saccharomyc
32	159.5	4.9	1128	1	BEM3_YEAST	P32873 saccharomyc
33	142	4.4	728	1	P85B_HUMAN	O00459 homo sapien
34	137.5	4.2	3093	1	POLG_BSTV1	Q65730 b genome po
35	137	4.2	724	1	P85B_BOVIN	P23726 bos taurus
36	135	4.2	1395	1	CUT1_MOUSE	P53564 mus musculu
37	135	4.2	3678	1	DMD_MOUSE	P11531 mus musculu
38	134	4.1	988	1	PCK1_SCHPO	P36582 schizosacch
39	132.5	4.1	724	1	P85A_BOVIN	P23727 bos taurus
40	132	4.1	705	1	YNP9_CAEEL	P34562 caenorhabdi
41	130.5	4.0	1080	1	HDA4_CHICK	P83038 gallus gall
42	130.5	4.0	3685	1	DMD_HUMAN	P11532 homo sapien
43	129	4.0	1505	1	CUT1_HUMAN	P39880 homo sapien
44	128.5	4.0	878	1	KPCO_HUMAN	Q9bz16 homo sapien
45	128.5	4.0	1616	1	P200_MYCGE	Q49429 mycoplasma

ALIGNMENTS

RESULT 1
RN_DROME
ID RN_DROME STANDARD; PRT; 384 AA.
AC P40809; Q9VI91;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE GTPase activating protein rnRacGAP.
GN RNRACGAP OR CG2595.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=Oregon-R; TISSUE=Pupae;
RX MEDLINE=93024458; PubMed=1406685;
RA Agnel M., Roeder L., Vola C., Griffin-Shea R.;
RT "A Drosophila rotund transcript expressed during spermatogenesis and
RT imaginal disc morphogenesis encodes a protein which is similar to
RT human Rac GTPase-activating (racGAP) proteins.";
RL Mol. Cell. Biol. 12:5111-5122(1992).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=96194886; PubMed=8654933;
RA Hoemann C.D., Bergeret E., Guichard A., Griffin-Shea R.;
RT "Alternative splicing of the Drosophila melanogaster rotundRacGAP
RT gene.";
RL Gene 168:135-141(1996).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

```
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: INVOLVED IN THE MORPHOGENESIS OF THE ADULT APPENDAGES.
CC       GTPASE-ACTIVATING PROTEIN FOR P21-RAC. PROMOTES THE EXCHANGE OF
CC       RAC-BOUND GDP BY GTP.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: IN PUPAE, EXPRESSED IN IMAGINAL DISKS AND ONLY
CC       IN THE MALE GONAD. IN ADULTS, ONLY FOUND IN THE TESTES, IN REGIONS
CC       FILLED WITH PRIMARY SPERMATOCYTES.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING SPERMATOGENESIS, IN PRIMARY
CC       SPERMATOCYTES, AND IMAGINAL DISK MORPHOGENESIS.
CC   -!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC       BINDING DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC   -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-23 IS THE INITIATOR.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M99702; AAA16480.2; -.
DR   EMBL; U22539; AAC47075.1; -.
DR   EMBL; U22539; AAC47076.1; -.
DR   EMBL; AE003672; AAF54034.1; -.
DR   PIR; A48122; A48122.
DR   FlyBase; FBgn0045843; rnRacGAP.
DR   InterPro; IPR002219; DAG_PE-bind.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00109; C1; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
KW   Guanine-nucleotide releasing factor; Spermatogenesis; Testis;
KW   Alternative splicing; Zinc; Phorbol-ester binding.
FT   DOMAIN       87    136       PHORBOL-ESTER AND DAG BINDING.
FT   DOMAIN      188    202       ARG/LYS-RICH.
FT   VARSPLIC    378    384       NLSSTHL -> RPMVSL (IN ISOFORM 2).
SQ   SEQUENCE   384 AA;  44217 MW;  D24959B833B7A2A8 CRC64;

  Query Match             14.7%;  Score 477;  DB 1;  Length 384;
  Best Local Similarity   40.7%;  Pred. No. 2e-23;
  Matches  110;  Conservative   44;  Mismatches  102;  Indels   14;  Gaps    6;

Qy        279 QSNGG-MRLHDFVSKT-VIKPESCVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCI 336
              ||: | :| |:|  |:       :|| | |||:|   ||:|| | :  |  |  :  : ||
Db         78 QSHSGLLREHNFKIKSYYYNVGNCVHCRKRIRFAMASLRCRACPLRCHIGCCRQLTVNCI 137

Qy        337 P-TLIGTPVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKE 395
             |   |||      | |:|:  : :||:|:::|||| ||| ||| : ||||:|      |
Db        138 PQPQIGTK----RGCLSDYAPRVAPMVPALIVHCVTEIEARGLQQEGLYRVSSTREKCKR 193

Qy        396 LKEKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSI 455
                |: | || |: | |    | | :|  :|||| | ||:     | | ||    |
Db        194 LRRKLLRGKSTPHLGN-KDTHTLCCCVKDFLRQLVHPLIPIYHRRDFEEATRHEDRLAVE 252

Qy        456 AAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPV 515
              |:| || || ||:|||||:||:| |::|:||  :| | ||| :| ||:       |
Db        253 MAVYLAVLELHQAHRDTLAYLMLHWQKIAESPAVRMTVNNLAVIFAPTLFG------DLD 306

Qy        516 TMSQDIKRQPKVVERLLSLPLEYWSQFMMV 545
                :::    :|:: || :|  :||||: |
Db        307 LTLENVVTWQRVLKVLLLMPAGFWSQFLEV 336


RESULT 2
MY9B_RAT
ID   MY9B_RAT       STANDARD;      PRT;  1980 AA.
AC   Q63358;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Myosin IXb (Unconventional myosin-9b).
GN   MYO9B OR MYR5.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain stem, and Spinal cord;
RX   MEDLINE=95188874; PubMed=7882973;
RA   Reinhard J., Scheel A.A., Diekmann D., Hall A., Ruppert C.,
RA   Baehler M.;
RT   "A novel type of myosin implicated in signalling by rho family
RT   GTPases.";
RL   EMBO J. 14:697-704(1995).
CC   -!- FUNCTION: MYOSINS ARE ACTIN-BASED MOTOR MOLECULES WITH ATPASE
CC       ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.
CC       MAY BE INVOLVED IN THE REMODELING OF THE ACTIN CYTOSKELETON. BINDS
CC       ACTIN WITH HIGH AFFINITY BOTH IN THE ABSENCE AND PRESENCE OF ATP
CC       AND ITS MECHANOCHEMICAL ACTIVITY IS INHIBITED BY CALCIUM IONS.
CC       ALSO ACTS AS A GTPASE ACTIVATING PROTEIN ON RHO.
CC   -!- SUBCELLULAR LOCATION: IN UNDIFFERENTIATED CELLS COLOCALIZES WITH
CC       F-ACTIN IN THE CELL PERIPHERY WHILE IN DIFFERENTIATED CELLS ITS
CC       LOCALIZTION IS CYTOPLASMIC WITH THE HIGHEST LEVELS IN THE
CC       PERINUCLEAR REGION (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN TESTIS, LUNG, THYMUS, BRAIN,
CC       LIVER, AND SPLEEN.
CC   -!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC       BINDING DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   -!- SIMILARITY: CONTAINS 4 IQ DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X77609; CAA54700.1; -.
DR   HSSP; P08799; 1MND.
DR   InterPro; IPR002219; DAG_PE-bind.
DR   InterPro; IPR000048; IQ_region.
DR   InterPro; IPR000159; RA_domain.
DR   InterPro; IPR000198; RhoGAP.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00063; myosin_head; 2.
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00612; IQ; 4.
DR   Pfam; PF00620; RhoGAP; 1.
DR   Pfam; PF00788; RA; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
```

```
DR   ProDom; PD000355; myosin_head; 2.
DR   SMART; SM00109; C1; 1.
DR   SMART; SM00015; IQ; 3.
DR   SMART; SM00242; MYSc; 1.
DR   SMART; SM00314; RA; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
DR   PROSITE; PS50096; IQ; 3.
KW   Myosin; Repeat; ATP-binding; Calmodulin-binding; Actin-binding;
KW   Cytoskeleton; Coiled coil; GTPase activation; Phorbol-ester binding;
KW   Zinc.
FT   DOMAIN        1    940       HEAD OR MOTOR DOMAIN.
FT   DOMAIN      941   1045       NECK OR REGULATORY DOMAIN.
FT   DOMAIN     1046   1980       TAIL.
FT   DOMAIN     1841   1861       COILED COIL (POTENTIAL).
FT   DOMAIN     1918   1948       COILED COIL (POTENTIAL).
FT   DOMAIN      845    856       ACTIN-BINDING.
FT   DOMAIN      958    978       IQ 1.
FT   DOMAIN      981   1001       IQ 2.
FT   DOMAIN     1002   1024       IQ 3.
FT   DOMAIN     1025   1054       IQ 4.
FT   DOMAIN     1593   1641       PHORBOL-ESTER AND DAG BINDING.
FT   DOMAIN     1673   1822       RHO-GAP.
FT   NP_BIND     239    246       ATP (POTENTIAL).
SQ   SEQUENCE   1980 AA;  225035 MW;  D79FEC4D0FAE0C05 CRC64;

  Query Match           10.7%;  Score 348;  DB 1;  Length 1980;
  Best Local Similarity   22.1%;  Pred. No. 3.3e-14;
  Matches  160;  Conservative  105;  Mismatches  262;  Indels  196;  Gaps   25;

Qy         82 EIKRRQRAEADCEKLERQIQLIR--EMLMCDTSGSIQLSEEQKSALAFLNRGQP---SSS 136
              :|:| |    | |:|   ::: |  ::     | |: |||: ::| | |   :    | |
Db       1272 QIQRYQ--HPDTERLATAVEIWRGKKLASAMLSQSLDLSEKPRTAGAALTPTEERRISFS 1329

Qy        137 NAGNKRLSTIDESGSILSDISFDK-------------------TDESLDWDSSLVKTFK 176
              :   :|| :  |  : :|:|  |                    |    |  |:  : |
Db       1330 TSDVSKLSPVKTSTEVDGDLSAKKPAGHKKKSEDPSAGPDAGLPTGSQGDSKSAFKRLFL 1389

Qy        177 LKKREKRRS------------------TSRQFVDGP--------------PGPVKKTRSI 204
               | ::|: |                  :|: :| |              |   : | :
Db       1390 HKAKDKKPSLEGVEETEGSGGQAAQEAPARKTLDVPSSQQHRHTTGEKPLKGKKNRNRKV 1449

Qy        205 G--SAVDQGNESIVAKTT------------VTVPND----GGPIEA------------VS 234
              |  :  ::  ||:  | |             :   ||      |||:            :
Db       1450 GQITVSEKWRESVFRKITNANELKFLDEFLLNKVNDLRSQKTPIESLFIEATERFRSNIK 1509

Qy        235 TIETVP-----------------------------------------------YWTRSRRKTGT 251
              |: :||                                              :|||  ||
Db       1510 TMYSVPNGKIHVGYKDLMENYQIVVSNLAAERGEKDTNLVLNVFQSLLDEFTRSYNKT-D 1568

Qy        252 LQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGK 311
               :|    :    |: |   :       : ||    | | |  |  |:|| |   |
Db       1569 FEPVKGKAQKKKRKQERAVQ-------EHNG----HVFASYQVNIPQSCEQCLSYIWLMD 1617

Qy        312 LSLKCRDCRVVSHPECRDRCPLPCIPT------LIGTPVKIGEGMLADFVSQTSPMIPSI 365
               :| |  |::  | :|  :    |  |      |   |      |:  | ::     :| :
Db       1618 KALLCSVCKMTCHKKCVHKIQSYCSYTGRRKSELGAEPGHF--GVCVDSLTSDKASVPIV 1675

Qy        366 VVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDD--IHAICSLLK 423
               :   :  :|  ||   |||| ||     :||::    ::| |   |::|  ||||  :||
Db       1676 LEKLLEHVEMHGLYTEGLYRKSGAANRTRELRQ---ALQTDPATVKLEDFPIHAITGVLK 1732

Qy        424 DFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRV 483
               :|| | |||:||     |: | |: ::   :||:|  :  ||:||  :|  |: || :|
Db       1733 QWLRELPEPLMTFAQYGDFLRAVELPEKQEQLAAIYAVLDHLPEANHTSLERLIFHLVKV 1792

Qy        484 A-QSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWSQF 542
              |    :|    || :| | ::     | ||:|  :|: :    || |:   :  :
Db       1793 ALLEDVNRMSPGALAIIFAPCLL-RCPDNSDPLTSMKDVLKITTCVEMLIKEQMRKYK-- 1849

Qy        543 MMVEQENIDPLHVIEN----------SNA-------FSTPQTPDIKVSLLGPVTTPEHQL 585
                |: | |: |   |:           ||       ||:|       |   ||     :|
```

```
Db       1850 --VKMEEINHLEAAESIAFRRLSLLRQNAPWPLKLGFSSPYEGVRTKSPRTPVVQDLEEL 1907

Qy        586 LKTPSSSS----------LSQRVRS------TLTKNTPRFGSKSKSATNLGRQGNFFASP 629
                |  ::          | :|::|       :|   |    :     ||  : :
Db       1908 GALPEEAAGGDEDREKEILMERIQSIKEEKEDITYRLPELDPRGSDEENLDSETSASTES 1967

Qy        630 MLK 632
               :|:
Db       1968 LLE 1970


RESULT 3
MY9B_MOUSE
ID   MY9B_MOUSE     STANDARD;      PRT;  2114 AA.
AC   Q9QY06; Q9QY07; Q9QY08; Q9QY09;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Myosin IXb (Unconventional myosin-9b).
GN   MYO9B OR MYR5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND VARIANTS.
RC   STRAIN=C57BL/6;
RX   MEDLINE=20047919; PubMed=10580159;
RA   Grewal P.K., Jones A.-M., Maconochie M., Lemmers R.J.F., Frants R.R.,
RA   Hewitt J.E.;
RT   "Cloning of the murine unconventional myosin gene Myo9b and
RT   identification of alternative splicing.";
RL   Gene 240:389-398(1999).
CC   -!- FUNCTION: MYOSINS ARE ACTIN-BASED MOTOR MOLECULES WITH ATPASE
CC       ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.
CC       MAY BE INVOLVED IN THE REMODELING OF THE ACTIN CYTOSKELETON. BINDS
CC       ACTIN WITH HIGH AFFINITY BOTH IN THE ABSENCE AND PRESENCE OF ATP
CC       AND ITS MECHANOCHEMICAL ACTIVITY IS INHIBITED BY CALCIUM IONS.
CC       ALSO ACTS AS A GTPASE ACTIVATING PROTEIN ON RHO.
CC   -!- SUBCELLULAR LOCATION: IN UNDIFFERENTIATED CELLS COLOCALIZES WITH
CC       F-ACTIN IN THE CELL PERIPHERY WHILE IN DIFFERENTIATED CELLS ITS
CC       LOCALIZTION IS CYTOPLASMIC WITH THE HIGHEST LEVELS IN THE
CC       PERINUCLEAR REGION (BY SIMILARITY).
CC   -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2/Q AND 3/C ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN THE BRAIN, HEART, MUSCLE AND
CC       INNER EAR.
CC   -!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC       BINDING DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   -!- SIMILARITY: CONTAINS 4 IQ DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF143687; AAF00122.1; -.
DR   EMBL; AF143685; AAF00120.1; -.
DR   EMBL; AF143686; AAF00121.1; ALT_FRAME.
DR   EMBL; AF143683; AAF00118.1; -.
DR   HSSP; P08799; 1MND.
DR   MGD; MGI:106624; Myo9b.
DR   InterPro; IPR002219; DAG_PE-bind.
DR   InterPro; IPR000048; IQ_region.
DR   InterPro; IPR000159; RA_domain.
DR   InterPro; IPR000198; RhoGAP.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00063; myosin_head; 2.
```

```
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00612; IQ; 4.
DR   Pfam; PF00620; RhoGAP; 1.
DR   Pfam; PF00788; RA; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 2.
DR   SMART; SM00109; C1; 1.
DR   SMART; SM00015; IQ; 4.
DR   SMART; SM00242; MYSc; 1.
DR   SMART; SM00314; RA; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
DR   PROSITE; PS50096; IQ; 3.
KW   Myosin; Repeat; ATP-binding; Calmodulin-binding; Actin-binding;
KW   Cytoskeleton; Coiled coil; GTPase activation; Phorbol-ester binding;
KW   Zinc; Alternative splicing; Polymorphism.
FT   DOMAIN        1    939       HEAD OR MOTOR DOMAIN.
FT   DOMAIN      940   1044       NECK OR REGULATORY DOMAIN.
FT   DOMAIN     1045   2114       TAIL.
FT   DOMAIN     1562   1592       COILED COIL (POTENTIAL).
FT   DOMAIN     1839   1859       COILED COIL (POTENTIAL).
FT   DOMAIN     1915   1945       COILED COIL (POTENTIAL).
FT   DOMAIN      844    855       ACTIN-BINDING.
FT   DOMAIN      957    977       IQ 1.
FT   DOMAIN      979   1000       IQ 2.
FT   DOMAIN     1001   1023       IQ 3.
FT   DOMAIN     1024   1053       IQ 4.
FT   DOMAIN     1591   1639       PHORBOL-ESTER AND DAG BINDING.
FT   DOMAIN     1671   1820       RHO-GAP.
FT   NP_BIND     239    246       ATP (POTENTIAL).
FT   VARSPLIC    740    740       R -> RCTGLDFSFERSEELDVNAFEDIMAFYESR (IN
FT                                ISOFORM 2).
FT   VARSPLIC   1875   1890       MISSING (IN ISOFORM 3).
FT   VARSPLIC   1977   2114       GPPAPALPCPISPTLSPLPEAAAPPRGRPTSFVTVRVKTPR
FT                                RTPIMPMANIKLPPGLPLHLTSWAPALQEAVVPVKRREPPA
FT                                RRQDQVHSVYIAPGADLPSQSTLIALDHDTILPGTKRRYSD
FT                                PPTYCLPPSSGQANG -> E (IN ISOFORM 3).
FT   VARIANT    1243   1243       L -> LEVSPVLPSSSL (ONLY IN STRAIN
FT                                C57B1/6; CONTAINS AN IN-FRAME 33 BP
FT                                IMPERFECT DUPLICATION).
FT   VARIANT    2083   2083       A -> P.
SQ   SEQUENCE   2114 AA;  238832 MW;  D774D4B1B2788045 CRC64;

  Query Match           10.5%;  Score 339.5;  DB 1;  Length 2114;
  Best Local Similarity   22.1%;  Pred. No. 1.3e-13;
  Matches  167;  Conservative  110;  Mismatches  276;  Indels  201;  Gaps   28;

Qy         50 ELGKYKDLLMKAETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIR--EML 107
              |  | ||   | |:   | |   |        :|:| |    | |:|   ::: |  ::
Db       1244 ESPKDKD---KDESSTKAQD---KPESPSGSTQIQRYQ--HPDTERLATAVEIWRGKKLA 1295

Qy        108 MCDTSGSIQLSEEQK---SALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDK---- 160
                | |: |||: :   :||      : | | :   :|| :  |  |  | |  |
Db       1296 SAVLSQSLDLSEKHRATGAALTPTEERRISFSTSDISKLSPVKTSAEIDGDFSSKKPSIH 1355

Qy        161 ----------TDESL------DWDSSLVKTFKLKKREKRRS----------------T 186
                       |  |      |  |:  : | |  | ::|: |
Db       1356 KKKSGDPSAGPDAGLSPGSQGDSKSAFKRLFLHKAKDKKPSLEGVEETESNGGQAAQETP 1415

Qy        187 SRQFVDGP--------------PGPVKKTRSIG--SAVDQGNESIVAKTT---------- 220
              :|: :| |              |   : | :|  :  ::  ||:  | |
Db       1416 ARKTLDVPSSQQHRHTTGEKPLKGKKNRNRKVGQITVSEKWRESVFRKITNANELKFLDE 1475

Qy        221 --VTVPND----GGPIEA------------VSTIETVP---------------------- 240
                :   ||      |||:             : |: :||
Db       1476 FLLNKVNDLRSQKTPIESLFIEATERFRSNIKTMYSVPNGKIHVGYKDLMENYQIVVSNL 1535

Qy        241 --------------------YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQ 279
                                 :|||  ||      : :   :  | : | :  :|   |
Db       1536 AAERGEKDTNLVLNVFQSLLDEFTRSYNKT------DFERAKSKAQKKKRKQERAV---Q 1586

Qy        280 SNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPT- 338
```

```
                : |   | | |  |  |:||  |   |     :| |  |::  | :|  :    |  |
Db       1587 EHNG---HVFASYQVNIPQSCEQCLSYIWLMDKALLCSVCKMTCHKKCVHKIQSYCSYTG 1643

Qy        339 -----LIGTPVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTV 393
                    |   |      |:  | ::     :| ::   :  :|  ||   |||| ||
Db       1644 RRKSELGAEPGHF--GVCVDSLTSDKASVPIVLEKLLEHVEMHGLYTEGLYRKSGAANRT 1701

Qy        394 KELKEKFLRVKTVPLLSKVDD--IHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDE 451
              :||::    ::| |   |::|  |||  :|| :|| | |||:||     |: | |: ::
Db       1702 RELRQ---ALQTDPAAVKLEDFPIHAITGVLKQWLRELPEPLMTFAQYGDFLRAVELPEK 1758

Qy        452 DNSIAAMYQAVGELPQANRDTLAFLMIHLQRVA-QSPHTKMDVANLAKVFGPTIVAHAVP 510
                ::|:|  :  ||:||  :|  |: || :||      :|    || :| | ::
Db       1759 QEQLSAIYAVLDHLPEANHTSLERLIFHLVKVALLEDVNRMSPGALAIIFAPCLL-RCPD 1817

Qy        511 NPDPVTMSQDIKRQPKVVERLLSLPLEYWSQFMMVEQENIDPLHVIEN----------SN 560
              | ||:|  :|: :    || |:    |   :: | : | |: |  |:          |
Db       1818 NSDPLTSMKDVLKITTCVEMLIK---EQMRKYKM-KMEEINHLEAAESIAFRRLSLLRQN 1873

Qy        561 A----------------FSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS---- 600
              |                  :|:|| ::   || ::       :      | :|::|
Db       1874 APWPLKLGFSSPYEGVRIKSPRTPVVQDLELGALSEEAAGGDEDREKEILMERIQSIKEE 1933

Qy        601 --TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
                :|   |    :     ||  : :    :|:
Db       1934 KEDITYRLPELDPRGSDEENLDSETSASTESLLE 1967


RESULT 4
CHIN_RAT
ID   CHIN_RAT       STANDARD;      PRT;   334 AA.
AC   P30337;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   N-chimaerin (NC) (N-chimerin) (Alpha chimerin) (A-chimaerin).
GN   CHN1 OR CHN.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Brain;
RX   MEDLINE=93074974; PubMed=1445199;
RA   Lim H.H., Michael G.J., Smith P., Lim L., Hall C.;
RT   "Developmental regulation and neuronal expression of the mRNA of rat
RT   n-chimaerin, a p21rac GAP:cDNA sequence.";
RL   Biochem. J. 287:415-422(1992).
CC   -!- FUNCTION: GTPASE ACTIVATING PROTEIN FOR P21-RAC AND A PHORBOL
CC       ESTER RECEPTOR. MAY PLAY AN IMPORTANT ROLE IN NEURONAL SIGNAL-
CC       TRANSDUCTION MECHANISMS.
CC   -!- TISSUE SPECIFICITY: IN NEURONS IN BRAIN REGIONS THAT ARE INVOLVED
CC       IN LEARNING AND MEMORY PROCESSES.
CC   -!- DEVELOPMENTAL STAGE: INCREASES IN AMOUNT DURING BRAIN DEVELOPMENT
CC       COINCIDENT WITH SYNAPTOGENESIS.
CC   -!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC       BINDING DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X67250; CAA47672.1; -.
DR   PIR; S25152; S25152.
DR   PIR; S29128; S29128.
DR   HSSP; P28867; 1PTQ.
DR   InterPro; IPR002219; DAG_PE-bind.
```

```
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   PRINTS; PR00008; DAGPEDOMAIN.
DR   SMART; SM00109; C1; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
KW   GTPase activation; Phorbol-ester binding; Zinc.
FT   DOMAIN       81    130       PHORBOL-ESTER AND DAG BINDING.
FT   DOMAIN      157    303       RHO-GAP.
SQ   SEQUENCE   334 AA;  38210 MW;  D7F828C63683528B CRC64;

  Query Match             10.1%;  Score 329;  DB 1;  Length 334;
  Best Local Similarity   32.4%;  Pred. No. 4.6e-14;
  Matches   91;  Conservative   46;  Mismatches  118;  Indels   26;  Gaps    8;

Qy        259 STLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLS--LKC 316
              :||   :  |: |              ::|:|   |   |  | |     :| ::  :||
Db         66 ATLKENEQIPKYE------------KVHNFKVHTFRGPHWCEYCA-NFMWGLIAQGVKC 111

Qy        317 RDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQR 376
               || :  | :|    |  | | |     |:    |   |       | :|  |: ||| |
Db        112 ADCGLNVHKQCSKMVPNDCKPDLKHVK-KVYSCDLTTLVKAHITKRPMVVDMCIREIESR 170

Qy        377 GLTETGLYRISGCDRTVKELKEKFLR--VKTVPLLSKVDDIHAICSLLKDFLRNLKEPLL 434
              ||   ||||:||    ::::|  | |   |    ::  :||: |   || : |:|  ||:
Db        171 GLNSEGLYRVSGFSDLIEDVKMAFDRDGEKADISVNMYEDINIITGALKLYFRDLPIPLI 230

Qy        435 TFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTK---M 491
              |:     |:|:|:| | |  :  :::|:  || |: :|| :|| ||:||  : | |   |
Db        231 TYDAYPKFIESAKIVDPDEQLETLHEALRSLPPAHCETLRYLMAHLKRV--TLHEKENLM 288

Qy        492 DVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLL 532
                ||  |||||::    |  ||:    ||: |  ||| |:
Db        289 SAENLGIVFGPTLMRS--PELDPMAALNDIRYQRLVVELLI 327


RESULT 5
CHIN_MOUSE
ID   CHIN_MOUSE     STANDARD;      PRT;   334 AA.
AC   Q91V57;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   N-chimaerin (NC) (N-chimerin) (Alpha chimerin) (A-chimaerin).
GN   CHN1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ILS, and ISS;
RA   Ehringer M.A., Thompson J., Conroy O., Xu Y., Yang F., Canniff J.,
RA   Beeson M., Gordon L., Bennett B., Johnson T.E., Sikela J.M.;
RT   "High-throughput sequence identification of gene coding variants
RT   within alcohol-related QTLs.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Colon, and Eye;
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: GTPASE ACTIVATING PROTEIN FOR P21-RAC AND A PHORBOL
CC       ESTER RECEPTOR. MAY PLAY AN IMPORTANT ROLE IN NEURONAL SIGNAL-
CC       TRANSDUCTION MECHANISMS (By similarity).
CC   -!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC       BINDING DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF332069; AAK56097.1; -.
DR   EMBL; AF332070; AAK56098.1; -.
DR   EMBL; BC010825; AAH10825.1; -.
DR   EMBL; BC024796; AAH24796.1; -.
DR   EMBL; BC025023; AAH25023.1; -.
DR   InterPro; IPR002219; DAG_PE-bind.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
KW   GTPase activation; Phorbol-ester binding; Zinc.
FT   DOMAIN       81    130       PHORBOL-ESTER AND DAG BINDING.
FT   DOMAIN      157    303       RHO-GAP.
SQ   SEQUENCE   334 AA;  38121 MW;  9EC5DA66D58A36C4 CRC64;

  Query Match             10.1%;  Score 328;  DB 1;  Length 334;
  Best Local Similarity   32.4%;  Pred. No. 5.4e-14;
  Matches   91;  Conservative   46;  Mismatches  118;  Indels   26;  Gaps    8;

Qy        259 STLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLS--LKC 316
              :||   :  |: |              ::|:|   |   |  | |     :| ::  :||
Db         66 ATLKENEQIPKYE------------KVHNFKVHTFRGPHWCEYCA-NFMWGLIAQGVKC 111

Qy        317 RDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQR 376
               || :  | :|    |  | | |     |:    |   |       | :|  |: ||| |
Db        112 ADCGLNVHKQCSKMVPNDCKPDLKHVK-KVYSCDLTTLVKAHITKRPMVVDMCIREIESR 170

Qy        377 GLTETGLYRISGCDRTVKELKEKFLR--VKTVPLLSKVDDIHAICSLLKDFLRNLKEPLL 434
              ||   ||||:||    ::::|  | |   |    ::  :||: |   || : |:|  ||:
Db        171 GLNSEGLYRVSGFSDLIEDVKMAFDRDGEKADISVNMYEDINIITGALKLYFRDLPIPLI 230

Qy        435 TFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTK---M 491
              |:     |:|:|:| | |  :  :::|:  || |: :|| :|| ||:||  : | |   |
Db        231 TYDAYPKFIESAKIMDPDEQLETLHEALRSLPPAHCETLRYLMAHLKRV--TLHEKENLM 288

Qy        492 DVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLL 532
                ||  |||||::    |  ||:    ||: |  ||| |:
Db        289 SAENLGIVFGPTLMRS--PELDPMAALNDIRYQRLVVELLI 327


RESULT 6
MY9B_HUMAN
ID   MY9B_HUMAN     STANDARD;      PRT;  2158 AA.
AC   Q13459; Q9UHN0; O75314; Q9NUJ2;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Myosin IXb (Unconventional myosin-9b).
GN   MYO9B OR MYR5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver, and Small intestine;
RX   MEDLINE=97063843; PubMed=8907710;
RA   Wirth J.A., Jensen K.A., Post P.L., Bement W.M., Mooseker M.S.;
RT   "Human myosin-IXb, an unconventional myosin with a chimerin-like
RT   rho/rac GTPase-activating protein domain in its tail.";
RL   J. Cell Sci. 109:653-661(1996).
RN   [2]
RP   PARTIAL SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX   MEDLINE=20047919; PubMed=10580159;
RA   Grewal P.K., Jones A.-M., Maconochie M., Lemmers R.J.F., Frants R.R.,
RA   Hewitt J.E.;
```

```
RT   "Cloning of the murine unconventional myosin gene Myo9b and
RT   identification of alternative splicing.";
RL   Gene 240:389-398(1999).
RN   [3]
RP   SEQUENCE OF 1940-2158 FROM N.A. (LONG ISOFORM).
RX   MEDLINE=98158729; PubMed=9490638;
RA   Post P.L., Bokoch G.M., Mooseker M.S.;
RT   "Human myosin-IXb is a mechanochemically active motor and a GAP for
RT   rho.";
RL   J. Cell Sci. 111:941-950(1998).
RN   [4]
RP   SEQUENCE OF 1828-2023 FROM N.A. (SHORT ISOFORM).
RC   TISSUE=Placenta;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA   Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,
RA   Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,
RA   Masuho Y., Kanehori K.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MYOSINS ARE ACTIN-BASED MOTOR MOLECULES WITH ATPASE
CC       ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.
CC       MAY BE INVOLVED IN THE REMODELING OF THE ACTIN CYTOSKELETON. BINDS
CC       ACTIN WITH HIGH AFFINITY BOTH IN THE ABSENCE AND PRESENCE OF ATP
CC       AND ITS MECHANOCHEMICAL ACTIVITY IS INHIBITED BY CALCIUM IONS.
CC       ALSO ACTS AS A GTPASE ACTIVATING PROTEIN ON RHO.
CC   -!- SUBCELLULAR LOCATION: IN UNDIFFERENTIATED CELLS COLOCALIZES WITH
CC       F-ACTIN IN THE CELL PERIPHERY WHILE IN DIFFERENTIATED CELLS ITS
CC       LOCALIZTION IS CYTOPLASMIC WITH THE HIGHEST LEVELS IN THE
CC       PERINUCLEAR REGION.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC       SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN PERIPHERAL BLOOD
CC       LEUKOCYTES AND AT LOWER LEVELS, IN THYMUS, SPLEEN, TESTIS,
CC       PROSTATE, OVARY, BRAIN, SMALL INTESTINE, AND LUNG.
CC   -!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC       BINDING DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC   -!- SIMILARITY: CONTAINS 4 IQ DOMAINS.
CC   -!- CAUTION: THE C-TERMINAL SEQUENCE FROM AA 1917 ONWARDS FROM REF.1
CC       WAS PROBABLY A CHIMERA.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U42391; AAC50402.1; ALT_SEQ.
DR   EMBL; AF143684; AAF00119.1; -.
DR   EMBL; AF020267; AAC26597.1; -.
DR   EMBL; AK002201; BAA92132.1; ALT_INIT.
DR   HSSP; P08799; 1MND.
DR   Genew; HGNC:7609; MYO9B.
DR   MIM; 602129; -.
DR   InterPro; IPR002219; DAG_PE-bind.
DR   InterPro; IPR000048; IQ_region.
DR   InterPro; IPR000159; RA_domain.
DR   InterPro; IPR000198; RhoGAP.
DR   InterPro; IPR001609; myosin_head.
DR   Pfam; PF00063; myosin_head; 2.
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00612; IQ; 4.
DR   Pfam; PF00620; RhoGAP; 1.
DR   Pfam; PF00788; RA; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 2.
DR   SMART; SM00109; C1; 1.
DR   SMART; SM00015; IQ; 3.
DR   SMART; SM00242; MYSc; 1.
DR   SMART; SM00314; RA; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
DR   PROSITE; PS50096; IQ; 3.
KW   Myosin; Repeat; ATP-binding; Calmodulin-binding; Actin-binding;
KW   Cytoskeleton; Coiled coil; GTPase activation; Phorbol-ester binding;
KW   Zinc; Alternative splicing.
FT   DOMAIN        1    939       HEAD OR MOTOR DOMAIN.
FT   DOMAIN      940   1044       NECK OR REGULATORY DOMAIN.
FT   DOMAIN     1045   2158       TAIL.
FT   DOMAIN     1046   1071       COILED COIL (POTENTIAL).
FT   DOMAIN     1880   1901       COILED COIL (POTENTIAL).
FT   DOMAIN     1959   1989       COILED COIL (POTENTIAL).
FT   DOMAIN      844    855       ACTIN-BINDING.
FT   DOMAIN      957    977       IQ 1.
FT   DOMAIN      979   1000       IQ 2.
FT   DOMAIN     1001   1023       IQ 3.
FT   DOMAIN     1024   1053       IQ 4.
FT   DOMAIN     1633   1681       PHORBOL-ESTER AND DAG BINDING.
FT   DOMAIN     1713   1862       RHO-GAP.
FT   NP_BIND     239    246       ATP (POTENTIAL).
FT   VARSPLIC   2022   2023       PP -> QY (IN SHORT ISOFORM).
FT   VARSPLIC   2024   2158       MISSING (IN SHORT ISOFORM).
FT   CONFLICT   1937   1939       QVP -> KT (IN REF. 4).
FT   CONFLICT   1947   1947       L -> P (IN REF. 4).
FT   CONFLICT   2040   2045       TVAAPP -> PWPPLH (IN REF. 3).
FT   CONFLICT   2049   2049       P -> L (IN REF. 3).
FT   CONFLICT   2067   2067       P -> S (IN REF. 3).
FT   CONFLICT   2157   2158       NG -> MAESHS (IN REF. 3).
SQ   SEQUENCE   2158 AA;  243556 MW;  4978F1D770F56D28 CRC64;

  Query Match             10.0%;  Score 323.5;  DB 1;  Length 2158;
  Best Local Similarity   22.3%;  Pred. No. 1.3e-12;
  Matches  149;  Conservative  117;  Mismatches  284;  Indels  119;  Gaps   27;

Qy         22 LSEGNEVQ--------FIQLAKDFEDFRKKWQ---RTDHELGKYKDLLMKAETERSALDV 70.
              || |::|          |:   ||     ||:      : |      ::::| | :  |:
Db       1404 LSPGSQVDSKSTFKRLFLHKTKD-----KKYSLEGAEELENAVSGHVVLEATTMKKGLEA 1458

Qy         71 ----KLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEEQKSALA 126
                  : :||  :      :| :      |  | :::          | | :||: : ::   .
Db       1459 PSGQQHRHAAGE------KRTKEPGGKGKKNRNVKI----------GKITVSEKWRESVF 1502

Qy        127 FLNRGQPSSSNAGNKRLSTIDE-SGSILSDISFDKTDESLDWDSSLVKTFKLKKREKRRS 185
                 | :::|     |  :||   : ::|:   ||          :::  ::  || ||
Db       1503 ----RQITNAN----ELKYLDEFLLNKINDLRSQKTP---------IESLFIEATEKFRS 1545

Qy        186 TSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIET-VPYWTR 244
               :    | | :    :|      |  ||     |          ::  :: :  :||
Db       1546 NIKTMYSVPNGKI----HVGYKDLMENYQIVVSNLATERGQKDTNLVLNLFQSLLDEFTR 1601

Qy        245 SRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCG 304
                 |       |    :   : : :         : ||    | | |  |  |:||  |
Db       1602 GYTK-------NDFEPVKQSKAQKKKRKQERAVQEHNG----HVFASYQVSIPQSCEQCL 1650

Qy        305 KRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTL--IGTP-VKIGE-GMLADFVSQTSP 360
                 |     :| |  |::  | :|  :    |  |    | |  : |  |:  | ::
Db       1651 SYIWLMDKALLCSVCKMTCHKKCVHKIQSHCSYTYGRKGEPGAEPGHFGVCVDSLTSDKA 1710

Qy        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDD--IHAI 418
               :| ::   :  :|  ||   |||| ||     :||::    ::| |   |:::  ||||
Db       1711 SVPIVLEKLLEHVEMHGLYTEGLYRKSGAANRTRELRQ---ALQTDPAAVKLENFPIHAI 1767

Qy        419 CSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMI 478
                :|| :|| | |||:||     |: | |: ::   :||:|  :  ||:|| ::|  |:
Db       1768 TGVLKQWLRELPEPLMTFAQYGDFLRAVELPEKQEQLAAIYAVLEHLPEANHNSLERLIF 1827

Qy        479 HLQRVA-QSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLE 537
               || :||      :|    || :| | ::     | ||:|  :|: :    || |:   :
Db       1828 HLVKVALLEDVNRMSPGALAIIFAPCLL-RCPDNSDPLTSMKDVLKITTCVEMLIKEQMR 1886

Qy        538 YWSQFMMVEQENIDPLHVIEN----------SNA-------FSTPQTPDIKVSLLGPVTT 580
```

```
              :    |: | |  |   |:          ||       ||:|    :  |   |
Db       1887 KYK----VKMEEISQLEAAESIAFRRLSLLRQNAPWPLKLGFSSPYEGVLNKSPQVPRDI 1942

Qy        581 PEHQ---LLKTPSSSS--------LSQRVRS------TLTKNTPRFGSKSKSATNLGRQG 623
               | :   ||:  ::         | :|::|       :|   |    :     ||  :
Db       1943 QEEELEVLLEEEAAGGDEDREKEILIERIQSIKEEKEDITYRLPELDPRGSDEENLDSET 2002

Qy        624 NFFASPMLK 632
              :    :|:
Db       2003 SASTESLLE 2011


RESULT 7
CHIN_HUMAN
ID   CHIN_HUMAN     STANDARD;      PRT;   459 AA.
AC   P15882; Q96FB0;
DT   01-APR-1990 (Rel. 14, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   N-chimaerin (NC) (N-chimerin) (Alpha chimerin) (A-chimaerin).
GN   CHN1 OR CHN OR ARHGAP2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM ALPHA-1).
RC   TISSUE=Retina;
RX   MEDLINE=90133942; PubMed=2299665;
RA   Hall C., Monfries C., Smith P., Lim H.H., Kozma R., Ahmed S.,
RA   Vanniasingham V., Leung T., Lim L.;
RT   "Novel human brain cDNA encoding a 34,000 Mr protein n-chimaerin,
RT   related to both the regulatory domain of protein kinase C and BCR,
RT   the product of the breakpoint cluster region gene.";
RL   J. Mol. Biol. 211:11-16(1990).
RN   [2]
RP   REVISIONS.
RX   MEDLINE=93074974; PubMed=1445199;
RA   Lim H.H., Michael G.J., Smith P., Lim L., Hall C.;
RT   "Developmental regulation and neuronal expression of the mRNA of rat
RT   n-chimaerin, a p21rac GAP:cDNA sequence.";
RL   Biochem. J. 287:415-422(1992).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM ALPHA-2).
RX   MEDLINE=93330292; PubMed=8336731;
RA   Hall C., Sin W.C., Teo M., Michael G.J., Smith P., Dong J.M.,
RA   Lim H.H., Manser E., Spurr N.K., Jones T.A., Lim L.;
RT   "Alpha 2-chimerin, an SH2-containing GTPase-activating protein for
RT   the ras-related protein p21rac derived by alternate splicing of the
RT   human n-chimerin gene, is selectively expressed in brain regions and
RT   testes.";
RL   Mol. Cell. Biol. 13:4986-4998(1993).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM ALPHA-2).
RC   TISSUE=Eye;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   PARTIAL SEQUENCE FROM N.A. (ISOFORM ALPHA-1).
RC   TISSUE=Fibroblast;
RX   MEDLINE=95172046; PubMed=7867622;
RA   Dong J.M., Smith P., Hall C., Lim L.;
RT   "Promoter region of the transcriptional unit for human alpha 1-
RT   chimaerin, a neuron-specific GTPase-activating protein for p21rac.";
RL   Eur. J. Biochem. 227:636-646(1995).
RN   [6]
RP   PHORBOL-ESTER BINDING.
RX   MEDLINE=91097509; PubMed=2268301;
RA   Ahmed S., Kozma R., Monfries C., Hall C., Lim H.H., Smith P., Lim L.;
RT   "Human brain n-chimaerin cDNA encodes a novel phorbol ester
RT   receptor.";
RL   Biochem. J. 272:767-773(1990).
CC   -!- FUNCTION: GTPASE ACTIVATING PROTEIN FOR P21-RAC AND A PHORBOL
CC       ESTER RECEPTOR. MAY PLAY AN IMPORTANT ROLE IN NEURONAL SIGNAL-
CC       TRANSDUCTION MECHANISMS.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA-1 AND ALPHA-2 (SHOWN
CC       HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: IN NEURONS IN BRAIN REGIONS THAT ARE INVOLVED
CC       IN LEARNING AND MEMORY PROCESSES.
CC   -!- DEVELOPMENTAL STAGE: INCREASES IN AMOUNT DURING BRAIN DEVELOPMENT
CC       COINCIDENT WITH SYNAPTOGENESIS.
CC   -!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC       BINDING DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X51408; CAA35769.1; ALT_INIT.
DR   EMBL; Z22641; CAA80354.1; -.
DR   EMBL; BC011393; AAH11393.1; -.
DR   EMBL; S75654; AAB33506.1; -.
DR   PIR; S08242; S08242.
DR   HSSP; P28867; 1PTQ.
DR   Genew; HGNC:1943; CHN1.
DR   MIM; 118423; -.
DR   InterPro; IPR002219; DAG_PE-bind.
DR   InterPro; IPR000198; RhoGAP.
DR   InterPro; IPR000980; SH2.
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   Pfam; PF00017; SH2; 1.
DR   PRINTS; PR00008; DAGPEDOMAIN.
DR   SMART; SM00109; C1; 1.
DR   ProDom; PD000093; SH2; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   GTPase activation; Phorbol-ester binding; Zinc; Alternative splicing;
KW   SH2 domain.
FT   DOMAIN       49    135       SH2.
FT   DOMAIN      206    255       PHORBOL-ESTER AND DAG BINDING.
FT   DOMAIN      282    428       RHO-GAP.
FT   VARSPLIC      1    183       MALTLFDTDEYRPPVWKSYLYQLQQEAPHPRRITCTCEVEN
FT                                RPKYYGREFHGMISREAADQLLIVAEGSYLIRESQRQPGTY
FT                                TLALRFGSQTRNFRLYYDGKHFVGEKRFESIHDLVTDGLIT
FT                                LYIETKAAEYIAKMTINPIYEHVGYTTLNREPAYKKHMPVL
FT                                KETHDERDSTGQDGVSEKR -> MPSKESWSGRKTNRAAVH
FT                                KSKQEGRQQDLLIAALGMKLGSPKSSVTIWQPLKLFAYSQ
FT                                (IN ISOFORM ALPHA-1).
SQ   SEQUENCE   459 AA;  53172 MW;  04C4CC9BCC611389 CRC64;

  Query Match             9.9%;  Score 322;  DB 1;  Length 459;
  Best Local Similarity   32.0%;  Pred. No. 2e-13;
  Matches   90;  Conservative   48;  Mismatches  117;  Indels   26;  Gaps    8;

Qy        259 STLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLS--LKC 316
              :||   :  |: |             ::|:|   |   | | |     :| ::  :||
Db        191 ATLKENEQIPKYE------------KIHNFKVHTFRGPHWCEYCA-NFMWGLIAQGVKC 236

Qy        317 RDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQR 376
               || :  | :|    |  | | |     |:    |   |   :   | :|  |: ||| |
Db        237 ADCGLNVHKQCSKMVPNDCKPDLKHVK-KVYSCDLTTLVKAHTTKRPMVVDMCIREIESR 295

Qy        377 GLTETGLYRISGCDRTVKELKEKFLR--VKTVPLLSKVDDIHAICSLLKDFLRNLKEPLL 434
              ||   ||||:||    ::::|  | |   |    ::  :||: |   || : |:|  ||:
Db        296 GLNSEGLYRVSGFSDLIEDVKMAFDRDGEKADISVNMYEDINIITGALKLYFRDLPIPLI 355

Qy        435 TFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTK---M 491
              |:     |:|:|:| | |  :  :::|:  || |: :|| :|| ||:||  : | |   |
```

```
Db        356 TYDAYPKFIESAKIMDPDEQLETLHEALKLLPPAHCETLRYLMAHLKRV--TLHEKENLM 413

Qy        492 DVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLL 532
              :  ||  |||||::    |  | :    ||: |  ||| |:
Db        414 NAENLGIVFGPTLMRS--PELDAMAALNDIRYQRLVVELLI 452


RESULT 8
CHIO_RAT
ID   CHIO_RAT       STANDARD;      PRT;   295 AA.
AC   Q03070;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Beta-chimaerin (Beta-chimerin).
GN   CHN2 OR BCH.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Testis;
RX   MEDLINE=93179371; PubMed=8440677;
RA   Leung T., How B.E., Manser E., Lim L.;
RT   "Germ cell beta-chimaerin, a new GTPase-activating protein for
RT   p21rac, is specifically expressed during the acrosomal assembly stage
RT   in rat testis.";
RL   J. Biol. Chem. 268:3813-3816(1993).
CC   -!- FUNCTION: GTPASE ACTIVATING PROTEIN FOR P21-RAC.
CC   -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; BETA-1 (SHOWN HERE) AND BETA-2;
CC       ARE PRODUCED BY ALTERNATIVE SPLICING. THE CEREBELLAR BETA-2
CC       FORM CONTAINS A SH2 DOMAIN.
CC   -!- TISSUE SPECIFICITY: FOUND IN CEREBELLUM AND TESTIS.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED SPECIFICALLY IN LATE STAGE
CC       SPERMATOCYTES. IN THE CEREBELLUM, EMERGENCE OF BETA-2 ISOFORM
CC       COINCIDES WITH GRANULE CELLS MATURATION AND EXHIBITS POSTNATAL
CC       DEVELOPMENTAL INCREASES. EXPRESSION IS SPECIFICALLY REDUCED IN
CC       WEAVER MUTANT.
CC   -!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC       BINDING DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X69489; CAA49244.1; -.
DR   EMBL; L07494; AAA40809.1; -.
DR   HSSP; P28867; 1PTQ.
DR   InterPro; IPR002219; DAG_PE-bind.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   PRINTS; PR00008; DAGPEDOMAIN.
DR   SMART; SM00109; C1; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
KW   GTPase activation; Phorbol-ester binding; Zinc; Membrane; SH2 domain;
KW   Alternative splicing.
FT   DOMAIN       42     91       PHORBOL-ESTER AND DAG BINDING.
FT   DOMAIN      118    262       RHO-GAP.
SQ   SEQUENCE   295 AA;  33837 MW;  D7692D957B4816BD CRC64;

  Query Match             9.1%;  Score 294;  DB 1;  Length 295;
  Best Local Similarity   29.3%;  Pred. No. 6.7e-12;
  Matches   81;  Conservative   52;  Mismatches  111;  Indels   32;  Gaps    8;

Qy        285 RLHDFVSKTVIKPESCVPCGKRIKFGKLS--LKCRDCRVVSHPECRDRCPLPCIPTLIGT 342
              : |:|   |   |  |  |     :| ::  ::| || :  | :|    |  | | |
Db         40 KTHNFKVHTFRGPHWCEYCA-NFMWGLIAQGVRCSDCGLNVHKQCSKHVPNDCQPDLKRI 98

Qy        343 PVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLR 402
               |:    |   |   :   | :|  |: ||| |||   ||||:||    ::::|  | |
Db         99 K-KVYCCDLTTLVKAHNTQRPMVVDICIREIEARGLKSEGLYRVSGFTEHIEDVKMAFDR 157

Qy        403 VKTVPLLSKVD-------DIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSI 455
                  | |       ||: |   || : |:|  |::|:     |:|||:|:: |  :
Db        158 DG-----EKADISANIYPDINIITGALKLYFRDLPIPIITYDTYTKFIEAAKISNADERL 212

Qy        456 AAMYQAVGELPQANRDTLAFLMIHLQRVAQSPHTK-MDVANLAKVFGPTIVAHAVPNPDP 514
              |::: :  || |: :|| :|||||::|  :     |:  ||  |||||::    |
Db        213 EAVHEVLMLLPPAHYETLRYLMIHLKKVTMNEKDNLMNAENLGIVFGPTLM--RPPEDST 270

Qy        515 VTMSQDIKRQPKVVERLLSLPLEYWSQFMMVEQENI 550
              :|   |:: |  :|:             :::| |::
Db        271 LTTLHDMRYQKLIVQ-------------ILIENEDV 293


RESULT 9
CHIO_HUMAN
ID   CHIO_HUMAN     STANDARD;      PRT;   468 AA.
AC   P52757;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Beta-chimaerin (Beta-chimerin).
GN   CHN2 OR BCH.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cerebellum;
RX   MEDLINE=94230370; PubMed=8175705;
RA   Leung T., How B.-E., Manser E., Lim L.;
RT   "Cerebellar beta 2-chimaerin, a GTPase-activating protein for p21
RT   ras-related rac is specifically expressed in granule cells and has a
RT   unique N-terminal SH2 domain.";
RL   J. Biol. Chem. 269:12888-12892(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RX   MEDLINE=95339337; PubMed=7614486;
RA   Yuan S., Miller D.W., Barnett G.H., Hahn J.F., Williams B.R.G.;
RT   "Identification and characterization of human beta 2-chimaerin:
RT   association with malignant transformation in astrocytoma.";
RL   Cancer Res. 55:3456-3461(1995).
RN   [3]
RP   SEQUENCE OF 18-192 FROM N.A.
RA   Strong C., Graves T., Yoakum M., Hawkins M.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: GTPASE ACTIVATING PROTEIN FOR P21-RAC. INSUFFICIENT
CC       EXPRESSION OF BETA-2 CHIMAERIN IS EXPECTED TO LEAD TO HIGHER RAC
CC       ACTIVITY AND COULD THEREFORE PLAY A ROLE IN THE PROGRESSION FROM
CC       LOW-GRADE TO HIGH-GRADE TUMORS.
CC   -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; BETA-1 AND BETA-2 (SHOWN HERE);
CC       ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: HIGEST LEVELS IN THE BRAIN AND PANCREAS. ALSO
CC       EXPRESSED IN THE HEART, PLACENTA, AND WEAKLY IN THE KIDNEY AND
CC       LIVER. EXPRESSION IS MUCH REDUCED IN THE MALIGNANT GLIOMAS,
CC       COMPARED TO NORMAL BRAIN OR LOW-GRADE ASTROCYTOMAS.
CC   -!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC       BINDING DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L29126; AAA19191.1; -.
DR   EMBL; U07223; AAA16836.1; -.
DR   EMBL; U28926; AAA86528.1; -.
DR   EMBL; AC004417; AAC06177.1; -.
DR   HSSP; P28867; 1PTQ.
DR   Genew; HGNC:1944; CHN2.
DR   MIM; 602857; -.
DR   InterPro; IPR002219; DAG_PE-bind.
DR   InterPro; IPR000198; RhoGAP.
DR   InterPro; IPR000980; SH2.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF00130; DAG_PE-bind; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   PRINTS; PR00008; DAGPEDOMAIN.
DR   ProDom; PD000093; SH2; 1.
DR   SMART; SM00109; C1; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   GTPase activation; Phorbol-ester binding; Zinc; SH2 domain;
KW   Alternative splicing.
FT   DOMAIN       59    127       SH2.
FT   DOMAIN      215    264       PHORBOL-ESTER AND DAG BINDING.
FT   DOMAIN      291    435       RHO-GAP.
FT   CONFLICT      1      6       MAASSN -> MRLL (IN REF. 1).
SQ   SEQUENCE   468 AA;  53923 MW;  63254958E0B5804C CRC64;

  Query Match             9.1%;  Score 293.5;  DB 1;  Length 468;
  Best Local Similarity   25.7%;  Pred. No. 1.4e-11;
  Matches   98;  Conservative   66;  Mismatches  155;  Indels   63;  Gaps   12;

Qy        199 KKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP------YWTRSRRKTGTL 252
              |:  ||   |  |  ::  :|           | :| : | |      | |  | |
Db        118 KRFESIHDLVTDGLITLYIETKAA--------EYISKMTTNPIYEHIGYATLLREKV--- 166

Qy        253 QPWNSDSTLNSRQLEPR---------TETDSVGTPQSNGGM----------RLHDFVSKT 293
                  |:  : |||         |  : : :      :           : |:|   |
Db        167 -----SRRLSRSKNEPRKTNVTHEEHTAVEKISSLVRRAALTHNDNHFNYEKTHNFKVHT 221

Qy        294 VIKPESCVPCGKRIKFGKLS--LKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGML 351
              |  |  |     :| ::  ::| || :  | :|    |  | | |     |:    |
Db        222 FRGPHWCEYCA-NFMWGLIAQGVRCSDCGLNVHKQCSKHVPNDCQPDLKRIK-KVYCCDL 279

Qy        352 ADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLR--VKTVPLL 409
              |   :   | :|  |: ||| |||   ||||:||    ::::|  | |   |
Db        280 TTLVKAHNTQRPMVVDICIREIEARGLKSEGLYRVSGFTEHIEDVKMAFDRDGEKADISA 339

Qy        410 SKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQAN 469
              :  ||: |   || : |:|  |::|:     |::||:|:: |  : |::: :  || |:
Db        340 NVYPDINIITGALKLYFRDLPIPVITYDTYSKFIDAAKISNADERLEAVHEVLMLLPPAH 399

Qy        470 RDTLAFLMIHLQRVAQSPHTK-MDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVV 528
             :|| :|||||::|  :     |:  ||  |||||::    |    :|   |:: |  :|
Db        400 YETLRYLMIHLKKVTMNEKDNFMNAENLGIVFGPTLM--RPPEDSTLTTLHDMRYQKLIV 457

Qy        529 ERLLSLPLEYWSQFMMVEQENI 550
              :             :::| |::
Db        458 Q------------ILIENEDV 466


RESULT 10
OPHL_HUMAN
ID   OPHL_HUMAN     STANDARD;      PRT;   814 AA.
AC   Q9UNA1; O75117; Q9UJ00; Q9BYS6; Q9BYS7;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Oligophrenin-1 like protein (GTPase regulator associated with focal
DE   adhesion kinase).
GN   OPHN1L OR GRAF OR KIAA0621.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Placenta, Heart, and Liver;
RA   Xia J.H., Tang X.X., Yu K.P., Pan Q., Dai H.P.;
RT   "Molecular cloning of human oligophrenin-1 like (OPHN1L) gene,
RT   complete CDS.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2), DISEASE, AND VARIANT LEUKEMIA SER-417.
RX   MEDLINE=20381355; PubMed=10908648;
RA   Borkhardt A., Bojesen S., Haas O.A., Fuchs U., Bartelheimer D.,
RA   Loncarevic I.F., Bohle R.M., Harbott J., Repp R., Jaeger U.,
RA   Viehmann S., Henn T., Korth P., Scharr D., Lampert F.;
RT   "The human GRAF gene is fused to MLL in a unique t(5;11)(q31;q23) and
RT   both alleles are disrupted in three cases of myelodysplastic
RT   syndrome/acute myeloid leukemia with a deletion 5q.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:9168-9173(2000).
RN   [3]
RP   SEQUENCE OF 53-785 FROM N.A. (ISOFORMS 1 AND 2).
RA   Bojesen S.E., Link C., Borkhardt A.;
RT   "Genomic structure of the human GRAF gene.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 62-814 FROM N.A. (ISOFORM 1).
RC   TISSUE=Brain;
RX   MEDLINE=98403880; PubMed=9734811;
RA   Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,
RA   Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. X.
RT   The complete sequences of 100 new cDNA clones from brain which can
RT   code for large proteins in vitro.";
RL   DNA Res. 5:169-176(1998).
CC   -!- FUNCTION: GTPase activating protein for RhoA.
CC   -!- SUBUNIT: Binds to the C-terminal of ppl25(FAK).
CC   -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC       produced by alternative splicing.
CC   -!- DISEASE: A form of juvenile myelomonocytic leukemia is
CC       characterized by a chromosomal translocation t(5;11)(q31;q23) that
CC       involves OPHN1L and MLL.
CC   -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF141884; AAD39482.1; -.
DR   EMBL; Y10388; CAA71414.2; -.
DR   EMBL; AJ309466; CAC29145.2; -.
DR   EMBL; AJ309467; CAC29145.2; JOINED.
DR   EMBL; AJ309468; CAC29145.2; JOINED.
DR   EMBL; AJ309469; CAC29145.2; JOINED.
DR   EMBL; AJ309470; CAC29145.2; JOINED.
DR   EMBL; AJ309471; CAC29145.2; JOINED.
DR   EMBL; AJ309472; CAC29145.2; JOINED.
DR   EMBL; AJ309473; CAC29145.2; JOINED.
DR   EMBL; AJ309474; CAC29145.2; JOINED.
DR   EMBL; AJ309475; CAC29145.2; JOINED.
```

```
DR   EMBL; AJ309476; CAC29145.2; JOINED.
DR   EMBL; AJ309477; CAC29145.2; JOINED.
DR   EMBL; AJ309478; CAC29145.2; JOINED.
DR   EMBL; AJ309479; CAC29145.2; JOINED.
DR   EMBL; AJ309480; CAC29145.2; JOINED.
DR   EMBL; AJ309481; CAC29145.2; JOINED.
DR   EMBL; AJ309482; CAC29145.2; JOINED.
DR   EMBL; AJ309483; CAC29145.2; JOINED.
DR   EMBL; AJ309484; CAC29145.2; JOINED.
DR   EMBL; AJ309485; CAC29145.2; JOINED.
DR   EMBL; AJ309486; CAC29145.2; JOINED.
DR   EMBL; AJ309487; CAC29145.2; JOINED.
DR   EMBL; AJ309466; CAC29146.2; -.
DR   EMBL; AJ309467; CAC29146.2; JOINED.
DR   EMBL; AJ309468; CAC29146.2; JOINED.
DR   EMBL; AJ309469; CAC29146.2; JOINED.
DR   EMBL; AJ309470; CAC29146.2; JOINED.
DR   EMBL; AJ309471; CAC29146.2; JOINED.
DR   EMBL; AJ309472; CAC29146.2; JOINED.
DR   EMBL; AJ309473; CAC29146.2; JOINED.
DR   EMBL; AJ309474; CAC29146.2; JOINED.
DR   EMBL; AJ309475; CAC29146.2; JOINED.
DR   EMBL; AJ309476; CAC29146.2; JOINED.
DR   EMBL; AJ309477; CAC29146.2; JOINED.
DR   EMBL; AJ309478; CAC29146.2; JOINED.
DR   EMBL; AJ309479; CAC29146.2; JOINED.
DR   EMBL; AJ309480; CAC29146.2; JOINED.
DR   EMBL; AJ309481; CAC29146.2; JOINED.
DR   EMBL; AJ309482; CAC29146.2; JOINED.
DR   EMBL; AJ309483; CAC29146.2; JOINED.
DR   EMBL; AJ309484; CAC29146.2; JOINED.
DR   EMBL; AJ309485; CAC29146.2; JOINED.
DR   EMBL; AJ309487; CAC29146.2; JOINED.
DR   EMBL; AB014521; BAA31596.1; -.
DR   MIM; 605370; -.
DR   HSSP; P19174; 2HSP.
DR   InterPro; IPR001849; PH.
DR   InterPro; IPR000198; RhoGAP.
DR   InterPro; IPR001452; SH3.
DR   Pfam; PF00169; PH; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   Pfam; PF00018; SH3; 1.
DR   PRINTS; PR00452; SH3DOMAIN.
DR   ProDom; PD000066; SH3; 1.
DR   SMART; SM00233; PH; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   SMART; SM00326; SH3; 1.
DR   PROSITE; PS50003; PH_DOMAIN; 1.
DR   PROSITE; PS50002; SH3; 1.
KW   GTPase activation; SH3 domain; Alternative splicing;
KW   Disease mutation; Chromosomal translocation; Proto-oncogene.
FT   DOMAIN      265    369       PH.
FT   DOMAIN      390    538       RHO-GAP.
FT   DOMAIN      756    814       SH3.
FT   DOMAIN      584    701       SER-RICH.
FT   VARSPLIC    700    754       MISSING (IN ISOFORM 2).
FT   VARIANT     417    417       N -> S (IN LEUKEMIA).
FT                                /FTId=VAR_013623.
FT   CONFLICT    355    355       E -> G (IN REF. 2 AND 3).
SQ   SEQUENCE   814 AA;  92234 MW;  5C81DBDECB32B18A CRC64;

  Query Match             9.0%;  Score 290.5;  DB 1;  Length 814;
  Best Local Similarity   23.5%;  Pred. No. 4.5e-11;
  Matches  152;  Conservative  116;  Mismatches  235;  Indels  143;  Gaps   29;

Qy          8 VRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK--AETER 65
              :||| :: :| :|  ||        |    | |||:      ::|  |:   |   |||:
Db         94 LRNLEDERIRMIENASE-------VLITPLEKFRKE------QIGAAKEAKKKYDKETEK 140

Qy         66 SALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIRE---MLMCDTSGSIQLSEEQK 122
                :  ||    :::  |::   |:  :: : |: |:|:     :  :    :|  :|:|
Db        141 YC-GILEKH----LNLSSKKK---ESQLQEADSQVDLVRQHFYEVSLEYVFKVQEVQERK 192

Qy        123 ------SALAFL-------NRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDS 169
                     ||||       : |   : : |: :             ||   |    :
Db        193 MFEFVEPLLAFLQGLFTFYHHGYELAKDFGDFKTQL---------TISIQNTRNRFEGTR 243

Qy        170 SLVKTFKLKKRE---KRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPND 226
              | |::   | :|   : :: |   ::|    |:: |  |::          |   |   |
Db        244 SEVESLMKKMKENPLEHKTISPYTMEGYL-YVQEKRHFGTS--------WVKHYCTYQRD 294

Qy        227 GGPIEAVSTIETVPYWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRL 286
                  |  ||:  :|  | |       | ::  :    | :|||:         :
Db        295 S------KQITMVPFDQKSGGKGG------EDESVILKSC-TRRKTDSI--------EKR 333

Qy        287 HDFVSKTVIKPESCVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKI 346
                |  : | :|          :  :| ::  | |   :   :|
Db        334 FCFDVEAVDRPGVITMQALSEEDRRLWMEAMDGREPVYNSNKD--------------SQ 378

Qy        347 GEGMLADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTV 406
              ||      :|   :  ||:  |:: :| ||: | ||||| | :  |::|    :  ||
Db        379 SEG-----TAQLDSIGFSIIRKCIHAVETRGINEQGLYRIVGVNSRVQKLLSVLMDPKTA 433

Qy        407 PLLSKVD-------DIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMY 459
                |: :         :|  | | || :|| |  ||: ::  |:|::||:: :::: :: ::
Db        434 ---SETETDICAEWEIKTITSALKTYLRMLPGPLMMYQFQRSFIKAAKLENQESRVSEIH 490

Qy        460 QAVGELPQANRDTLAFLMIHLQRVAQS-PHTKMDVANLAKVFGPTIVAHAVPNPDPVTMS 518
              |  ||: ||  |  || ||  || :     | ||||  |||||::    |  : |
Db        491 SLVHRLPEKNRQMLQLLMNHLANVANNHKQNLMTVANLGVVFGPTLLR---PQEETVAAI 547

Qy        519 QDIKRQPKVVERLL-----------SLPLEYWSQFMMVEQENID---------PL---HV 555
              ||| |  |:| |:            :||   :|  :  ::: |          ||   |
Db        548 MDIKFQNIVIEILIENHEKIFNTVPDMPLTN-AQLHLSRKKSSDSKPPSCSERPLTLFHT 606

Qy        556 IENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRST 601
              ::::       |   |  | |  |::  : :|   |||||   : |:
Db        607 VQSTE--KQEQRNSIINSSLESVSSNPNSILN--SSSSLQPNMNSS 648
```

```
RESULT 11
ABR_HUMAN
ID   ABR_HUMAN      STANDARD;      PRT;   859 AA.
AC   Q12979; Q13693; Q13694;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Active breakpoint cluster region-related protein.
GN   ABR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Hippocampus;
RX   MEDLINE=94086546; PubMed=8262969;
RA   Tan E.-C., Leung T., Manser E., Lim L.;
RT   "The human active breakpoint cluster region-related gene encodes a
RT   brain protein with homology to guanine nucleotide exchange proteins
RT   and GTPase-activating proteins.";
RL   J. Biol. Chem. 268:27291-27298(1993).
RN   [2]
RP   SEQUENCE OF 39-859 FROM N.A. (LONG AND SHORT FORMS).
RC   TISSUE=Fibroblast;
RX   MEDLINE=93352461; PubMed=8349582;
RA   Heisterkamp N., Kaartinen V., van Soest S., Bokoch G.M., Groffen J.;
RT   "Human ABR encodes a protein with GAPrac activity and homology to the
RT   DBL nucleotide exchange factor domain.";
RL   J. Biol. Chem. 268:16903-16906(1993).
RN   [3]
RP   SEQUENCE OF 436-597 FROM N.A.
RX   MEDLINE=90067847; PubMed=2587217;
RA   Heisterkamp N., Morris C., Groffen J.;
RT   "ABR, an active BCR-related gene.";
RL   Nucleic Acids Res. 17:8821-8831(1989).
```

```
CC   -!- FUNCTION: GTPASE-ACTIVATING PROTEIN FOR RAC AND CDC42. PROMOTES
CC       THE EXCHANGE OF RAC OR CDC42-BOUND GDP BY GTP, THEREBY ACTIVATING
CC       THEM.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC       SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: HIGHLY ENRICHED IN THE BRAIN. MUCH WEAKER
CC       EXPRESSION IN HEART, LUNG AND MUSCLE.
CC   -!- SIMILARITY: CONTAINS 1 DBL-HOMOLOGY (DH) DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC   -!- SIMILARITY: STRONG, TO HUMAN BCR.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U01147; AAC50063.1; -.
DR   EMBL; L19704; AAC37519.1; -.
DR   EMBL; L19705; AAC37518.1; ALT_INIT.
DR   Genew; HGNC:81; ABR.
DR   MIM; 600365; -.
DR   InterPro; IPR000008; C2.
DR   InterPro; IPR001331; GDS_CDC24.
DR   InterPro; IPR001849; PH.
DR   InterPro; IPR000198; RhoGAP.
DR   InterPro; IPR000219; RhoGEF.
DR   Pfam; PF00168; C2; 1.
DR   Pfam; PF00169; PH; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   Pfam; PF00621; RhoGEF; 1.
DR   SMART; SM00239; C2; 1.
DR   SMART; SM00233; PH; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   SMART; SM00325; RhoGEF; 1.
DR   PROSITE; PS50004; C2_DOMAIN_2; 1.
DR   PROSITE; PS50010; DH_2; 1.
DR   PROSITE; PS00741; DH_1; 1.
DR   PROSITE; PS50003; PH_DOMAIN; 1.
KW   Guanine-nucleotide releasing factor; Alternative splicing.
FT   DOMAIN       91    284       DH.
FT   DOMAIN      301    459       PH.
FT   DOMAIN      463    595       C2 DOMAIN.
FT   DOMAIN      661    805       RHO-GAP.
FT   DOMAIN      417    420       POLY-LEU.
FT   VARSPLIC      1     82       MEPLSHRGLPRLSWIDTLYSNFSYGTDEYDGEGNEEQKGPP
FT                                EGSETMPYIDESPTMSPQLSARSQGRGDGVSPTPPEGLAPG
FT                                -> MEEEEEAIGLLDKVLEDEDVFLLEECELGTPTSPGSGS
FT                                PFLVAVK (IN SHORT ISOFORM).
FT   CONFLICT     67     67       R -> G (IN REF. 2).
FT   CONFLICT    657    660       RSKV -> VQGA (IN REF. 2).
FT   CONFLICT    761    761       L -> V (IN REF. 2).
SQ   SEQUENCE   859 AA;  97696 MW;  9FD50CD54FA99483 CRC64;

  Query Match             8.7%;  Score 282;  DB 1;  Length 859;
  Best Local Similarity   25.5%;  Pred. No. 1.7e-10;
  Matches  140;  Conservative   67;  Mismatches  201;  Indels  140;  Gaps   21;

Qy         48 DHELGKYKDLLMKAETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREML 107
              ||||   :|: ||    :| :         | :   | : ||    |:|::::    | |
Db        374 DHEL---EDMKMKISALKSEI---------QKEKANKGQSRA---IERLKKKM-FENEFL 417

Qy        108 MCDTSGSIQLSEEQKSALAFL----NRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDE 163
              :   | :|      ::  ::|    :  : |       ::|   |    :|| :
Db        418 LLLNSPTIPFRIHNRNGKSYLFLLSSDYERSEWREAIQKLQKKDLQAFVLSSVELQVLTG 477

Qy        164 SLDWDSSLVKTFKLKKREKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTV 223
              |          |||:       || :  | ||       |  :     :|     |: |
Db        478 S---------CFKLRTVHNIPVTSNKDDDESPGLYGFLHVIVHSAKGFKQSANLYCTLEV 528

Qy        224 PNDGGPIEAVSTIETVPYWTRSRRKT----GTLQP-WNS-------------------- 257
              : |             |:  |: ||     | :| |:
Db        529 DSFG-------------YFV-SKAKTRVFRDTAEPKWDEEFEIELEGSQSLRILCYEKCY 574

Qy        258 DSTLNSR---------------QLEPRT-ETDSVGTPQSNGGMRLHDFVSKTVIKPESCV 301
              | |  ::               ||:|:| || :  |          | :    || |  :
Db        575 DKTKVNKDNNEIVDKIMGKGQIQLDPQTVETKNWHT----------DVIEMNGIKVEFSM 624

Qy        302 PCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVS----Q 357
                  |     :|||                          || |   |:    :|    :
Db        625 KFTSR----DMSLK------------------------RTPSKKQTGVFGVKISVVTKR 655

Qy        358 TSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKF-LRVKTVPLLSKVDDIH 416
                :| ||  || |:|:||: | |:|||||    :: ||  |    | : |:    ||:
Db        656 ERSKVPYIVRQCVEEVEKRGIEEVGIYRISGVATDIQALKAVFDANNKDILLMLSDMDIN 715

Qy        417 AICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFL 476
              ||   || : | | ||||| ||  ||||   ::|       |   :  ||  |  |  ||
Db        716 AIAGTLKLYFRELPEPLLTDRLYPAFMEGIALSDPAAKENCMMHLLRSLPDPNLITFLFL 775

Qy        477 MIHLQRVAQ-SPHTKMDVANLAKVFGPTIV--------AHAVPNPDPVTMSQDIKRQPKV 527
              : ||:|||:  |  || : ||| |||||::         ||     |    | |:  | :|
Db        776 LEHLKRVAEKEPINKMSLHNLATVFGPTLLRPSEVESKAHLTSAAD--IWSHDVMAQVQV 833

Qy        528 VERLLSLP 535
              :   | |
Db        834 LLYYLQHP 841


RESULT 12
YN54_CAEEL
ID   YN54_CAEEL     STANDARD;      PRT;   747 AA.
AC   P34588;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Hypothetical protein ZC21.4 in chromosome III.
GN   ZC21.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA   Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA   Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA   Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA   Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA   Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA   Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
CC   -!- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L16685; AAA28171.1; ALT_INIT.
DR   HSSP; Q07960; 1RGP.
```

```
DR   WormPep; ZC21.4; CE00534.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00324; RhoGAP; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   747 AA;  80829 MW;  48E9D61EDC4C56B0 CRC64;

  Query Match             8.6%;  Score 280.5;  DB 1;  Length 747;
  Best Local Similarity   25.1%;  Pred. No. 1.7e-10;
  Matches  105;  Conservative   59;  Mismatches  146;  Indels  109;  Gaps   14;

Qy        223 VPNDGGPIEAVSTIETVPYWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNG 282
              : |   |    :| ::   | :|:       :| |:  |:|         : | ||
Db         35 IENGEAPTATATTPKSGRKWKKSKAAKQGSGGGSSGSSSGSQQ------QGAAGAPQPVL 88

Qy        283 GMRLHDFVSKTVIKPESCVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGT 342
              |:|: |            | |                        | |
Db         89 GVRIAD-----------CPTG--------------------SCEDH------------ 103

Qy        343 PVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEK--- 399
                              :| ||  ||  ||  |:   |:||| |    |  |||
Db        104 -------------------VPMIVQACVCVIETYGMDTVGIYRIPGNTAAVNALKESLSN 144

Qy        400 --FLRVKTVPLLSKVD-------DIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITD 450
                |  |     ||||:       |:: : |||| ||| | |||||  :|   |::|  |:
Db        145 RGFDSVD----LSKVESLDPRWRDVNVVSSLLKMFLRKLPEPLLTDKLYPFFIDANRIST 200

Qy        451 EDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQ-SPHTKMDVANLAKVFGPTIVAHAV 509
              | :  :   : :||: : ||| ||::||  : : |   ||:  ||| :|||:||
Db        201 HHNRLHKLRNLLRKLPRPHYDTLRFLIVHLSEITKHSDVNKMECRNLALMFGPSIVR--- 257

Qy        510 PNPDPV-TMSQDIKRQPKVVERLLSLPLEYWSQFMMVEQENIDPLHVIENSNAFSTPQTP 568
              |: | : ||   :  | |::| |:   |  : :    | : :   |  :  |    |  |
Db        258 PSDDNMATMVTHMSDQCKIIETLIHYNLWMFDESSTTE-DAVPEQHPADGQN----PLEP 312

Qy        569 -----DIKVSLLGPVTTPEHQLLKTPSSSSLS--------QRVRSTLTKNTPRFGSKSK 614
                   :   :        | |::  :    :        |:::: | :|: |  ||||
Db        313 GGYGVGVPTGVSAASFNDMHNLIRKANEDQAAAMMNEGKGQKIKNMLRRNSRRDKSKSK 371


RESULT 13
BCR_HUMAN
ID   BCR_HUMAN      STANDARD;      PRT;  1271 AA.
AC   P11274; Q12842;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1989 (Rel. 11, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Breakpoint cluster region protein (EC 2.7.1.-).
GN   BCR OR BCR1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95394474; PubMed=7665185;
RA   Chissoe S.L., Bodenteich A., Wang Y.-F., Wang Y.-P., Burian D.,
RA   Clifton S.W., Crabtree J., Freeman A., Iyer K., Jian L., Ma Y.,
RA   McLaury H.-J., Pan H.-Q., Sarhan O.H., Toth S., Wang Z., Zhang G.,
RA   Heisterkamp N., Groffen J., Roe B.A.;
RT   "Sequence and analysis of the human ABL gene, the BCR gene, and
RT   regions involved in the Philadelphia chromosomal translocation.";
RL   Genomics 27:67-82(1995).
RN   [2]
RP   SEQUENCE OF 1-872 FROM N.A.
RX   MEDLINE=87218455; PubMed=3107980;
RA   Hariharan I.K., Adams J.M.;
RT   "cDNA sequence for human bcr, the gene that translocates to the abl
RT   oncogene in chronic myeloid leukaemia.";
RL   EMBO J. 6:115-119(1987).
RN   [3]
RP   SEQUENCE OF 683-1271 FROM N.A.
RX   MEDLINE=85240564; PubMed=2989703;
RA   Heisterkamp N., Stam K., Groffen J., de Klein A., Grosveld G.;
RT   "Structural organization of the bcr gene and its role in the Ph'
RT   translocation.";
RL   Nature 315:758-761(1985).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88217290; PubMed=3285291;
RA   Lifshitz B., Fainstein E., Marcelle C., Shtivelman E., Amson R.,
RA   Gale R.P., Canaani E.;
RT   "bcr genes and transcripts.";
RL   Oncogene 2:113-117(1988).
RN   [5]
RP   SEQUENCE OF 1-693 FROM N.A.
RX   MEDLINE=87092329; PubMed=3540951;
RA   Mes-Masson A.M., McLaughlin J., Daley G.Q., Paskind M., Witte O.N.;
RT   "Overlapping cDNA clones define the complete coding region for the
RT   P210c-abl gene product associated with chronic myelogenous leukemia
RT   cells containing the Philadelphia chromosome.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:9768-9772(1986).
RN   [6]
RP   REVISIONS.
RA   Mes-Masson A.M., McLaughlin J., Daley G.Q., Paskind M., Witte O.N.;
RL   Proc. Natl. Acad. Sci. U.S.A. 84:2507-2507(1987).
RN   [7]
RP   SEQUENCE OF 1-46 AND 275-426 FROM N.A.
RX   MEDLINE=91088292; PubMed=2263470;
RA   Zhu Q.S., Heisterkamp N., Groffen J.;
RT   "Unique organization of the human BCR gene promoter.";
RL   Nucleic Acids Res. 18:7119-7125(1990).
RN   [8]
RP   SEQUENCE OF 56-426 FROM N.A.
RX   MEDLINE=88065859; PubMed=2825022;
RA   Fainstein E., Marcelle C., Rosner A., Canaani E., Gale R.P.,
RA   Dreazen O., Smith S.D., Croce C.M.;
RT   "A new fused transcript in Philadelphia chromosome positive acute
RT   lymphocytic leukaemia.";
RL   Nature 330:386-388(1987).
RN   [9]
RP   SEQUENCE OF 1-4 FROM N.A.
RX   MEDLINE=91172169; PubMed=1900918;
RA   Shah N.P., Witte O.N., Denny C.T.;
RT   "Characterization of the BCR promoter in Philadelphia chromosome-
RT   positive and -negative cell lines.";
RL   Mol. Cell. Biol. 11:1854-1860(1991).
RN   [10]
RP   FUNCTION.
RX   MEDLINE=91238969; PubMed=1903516;
RA   Diekmann D., Brill S., Garrett M.D., Totty N., Hsuan J.,
RA   Monfries C., Hall C., Lim L., Hall A.;
RT   "Bcr encodes a GTPase-activating protein for p21rac.";
RL   Nature 351:400-402(1991).
RN   [11]
RP   BINDING TO ABL SH2-DOMAIN.
RX   MEDLINE=91300547; PubMed=1712671;
RA   Pendergast A.M., Muller A.J., Havlik M.H., Maru Y., Witte O.N.;
RT   "BCR sequences essential for transformation by the BCR-ABL oncogene
RT   bind to the ABL SH2 regulatory domain in a non-phosphotyrosine-
RT   dependent manner.";
RL   Cell 66:161-171(1991).
RN   [12]
RP   KINASE ACTIVITY.
RX   MEDLINE=92034969; PubMed=1657398;
RA   Maru Y., Witte O.N.;
RT   "The BCR gene encodes a novel serine/threonine kinase activity within
RT   a single exon.";
RL   Cell 67:459-468(1991).
CC   -!- FUNCTION: GTPASE-ACTIVATING PROTEIN FOR RAC1 AND CDC42. PROMOTES
CC       THE EXCHANGE OF RAC OR CDC42-BOUND GDP BY GTP, THEREBY ACTIVATING
CC       THEM. DISPLAYS SERINE/THREONINE KINASE ACTIVITY.
CC   -!- DOMAIN: THE REGION INVOLVED IN BINDING TO ABL SH2-DOMAIN IS RICH
CC       IN SERINE RESIDUES AND NEEDS TO BE SER/THR PHOSPHORYLATED PRIOR TO
CC       SH2 BINDING. THIS REGION IS ESSENTIAL FOR THE ACTIVATION OF THE
CC       ABL TYROSINE KINASE AND TRANSFORMING POTENTIAL OF THE CHIMERIC
```

```
CC       BCR-ABL ONCOGENE.
CC   -!- PTM: AUTOPHOSPHORYLATED.
CC   -!- DISEASE: Participates in a t(9;22)(q34;q11) chromosomal
CC       translocation that produces a BCR-ABL oncogene responsible for
CC       chronic myeloid leukemia (CML), acute myeloid leukemia (AML) and
CC       acute lymphoblastic leukemia (ALL).
CC   -!- SIMILARITY: CONTAINS 1 DBL-HOMOLOGY (DH) DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC   -!- SIMILARITY: STRONG, TO HUMAN ABR AND DROSOPHILA ROTUND.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U07000; AAB60388.1; -.
DR   EMBL; M24603; AAA65939.1; -.
DR   EMBL; Y00661; CAA68676.1; -.
DR   EMBL; X02596; CAA26441.1; -.
DR   EMBL; M15025; AAA35594.1; -.
DR   EMBL; X52828; CAA37010.1; -.
DR   EMBL; X52829; CAA37011.1; -.
DR   EMBL; M64437; -; NOT_ANNOTATED_CDS.
DR   PIR; A26172; TVHUBR.
DR   PIR; A26664; TVHUA2.
DR   PIR; A28765; TVHUA3.
DR   Genew; HGNC:1014; BCR.
DR   MIM; 151410; -.
DR   InterPro; IPR000008; C2.
DR   InterPro; IPR001331; GDS_CDC24.
DR   InterPro; IPR001849; PH.
DR   InterPro; IPR000198; RhoGAP.
DR   InterPro; IPR000219; RhoGEF.
DR   Pfam; PF00168; C2; 1.
DR   Pfam; PF00169; PH; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   Pfam; PF00621; RhoGEF; 1.
DR   SMART; SM00239; C2; 1.
DR   SMART; SM00233; PH; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   SMART; SM00325; RhoGEF; 1.
DR   PROSITE; PS50004; C2_DOMAIN_2; 1.
DR   PROSITE; PS50010; DH_2; 1.
DR   PROSITE; PS00741; DH_1; 1.
DR   PROSITE; PS50003; PH_DOMAIN; 1.
KW   Guanine-nucleotide releasing factor; Proto-oncogene; Transferase;
KW   Chromosomal translocation; Serine/threonine-protein kinase;
KW   Phosphorylation.
FT   DOMAIN        1    426       KINASE.
FT   DOMAIN      498    691       DH.
FT   DOMAIN      708    866       PH.
FT   DOMAIN      870   1002       C2 DOMAIN.
FT   DOMAIN     1068   1212       RHO-GAP.
FT   BINDING     197    385       TO ABL SH2-DOMAIN.
FT   DOMAIN      824    827       POLY-LEU.
FT   SITE        426    427       BREAKPOINT FOR TRANSLOCATION TO FORM
FT                                BCR-ABL ONCOGENE.
FT   CONFLICT    287    287       M -> I (IN REF. 4).
FT   CONFLICT    418    418       G -> D (IN REF. 4).
FT   CONFLICT    483    483       E -> K (IN REF. 4).
FT   CONFLICT    560    560       F -> S (IN REF. 4).
FT   CONFLICT    733    733       E -> D (IN REF. 1 AND 4).
FT   CONFLICT    796    796       S -> N (IN REF. 1).
FT   CONFLICT    961   1004       MISSING (IN REF. 4).
SQ   SEQUENCE   1271 AA;  142805 MW;  A59F76B6F1AF3236 CRC64;

  Query Match             8.5%;  Score 275.5;  DB 1;  Length 1271;
  Best Local Similarity   24.0%;  Pred. No. 7.5e-10;
```

```
  Matches  137;  Conservative   86;  Mismatches  196;  Indels  153;  Gaps   28;

Qy        43 KW---------QRTDHELGKYKDLLMKAETERSALDVKLKHARNQVDVEIKRRQRAEADC 93
             ||         |  | ||    :: :  : |  || :|:    :|:  :|:| :||
Db       754 KWYIPLTDLSFQMVD-ELEAVPNIPLVPDEELDALKIKI----SQIKSDIQREKRANKGS 808

Qy        94 EKLERQIQLIREMLMCDTSGSIQLSEEQKSALAFLNRGQPSSSNAGNKRLSTIDESGSIL 153
             :  ||    ::: |            ||:| |  ::       :: | :  |      |
Db       809 KATER----LKKKL-----------SEQESLLLLMSPSMAFRVHSRNGKSYTF----LIS 849

Qy       154 SDISFDKTDESLDWDSSLVKTFKLKKREKRRSTSRQFVDGPPGPVKKTRSIGSAVDQ--G 211
             ||        |  :|  ::         ||:::   | |                |: |  |
Db       850 SDY------ERAEWRENI--------REQQKKCFRSF-----------SLTSVELQMLT 883

Qy       212 NESIVAKTTVTVP------NDGGP--------------------IEAVSTIETVPYW-- 242
             |  :  :|  ::|      :|  |                    :     :::  |:
Db       884 NSCVKLQTVHSIPLTINKEDDESPGLYGFLNVIVHSATGFKQSSNLYCTLEVDSFGYFVN 943

Qy       243 -TRSRRKTGTLQP-WNSDSTL---NSRQLE-------------PRTETDSVGTPQSNGGM 284
              .::|    | :| || :  :    |: |               |: : :|       | :
Db       944 KAKTRVYRDTAEPNWNEEFEIELEGSQTLRILCYEKCYNKTKIPKEDGESTDRLMGKGQV 1003

Qy       285 RL-------HDFVSKTVIKPESCVPCGKRIKFG--KLSLKCRDCRVVSHPECRDRCPLPC 335
             :|        |:  :|||     :     , :||   : |||             | |
Db      1004 QLDPQALQDRDW-QRTVIAMNG-IEVKLSVKFNSREFSLK-------------RMPSRK 1047

Qy       336 IPTLIGTPVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKE 395
               : |  |||     |    :    :| ||  || |||:||: | |:||:||    ::
Db      1048 QTGVFG--VKI-----AVVTKRERSKVPYIVRQCVEEIERRGMEEVGIYRVSGVATDIQA 1100

Qy       396 LKEKF-LRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNS 454
             ||  | :  | | ::    |::||   || : | | ||| |      | |   ::|      ,
Db      1101 LKAAFDVNNKDVSVMMSEMDVNAIAGTLKLYFRELPEPLFTDEFYPNFAEGIALSDPVAK 1160

Qy       455 IAAMYQAVGELPQANRDTLAFLMIHLQRVAQSPH-TKMDVANLAKVFGPTIV------AH 507
               : |  :  ||:||  |  ||: ||:|||:     || : ||| |||||::      :
Db      1161 ESCMLNLLLSLPEANLLTFLFLLDHLKRVAEKEAVNKMSLHNLATVFGPTLLRPSEKESK 1220

Qy       508 AVPNP-DPVTM----SQDIKRQPKVVERLLSL 534
               ||  |:||    | ::  | :|:   | |
Db      1221 LPANPSQPITMTDSWSLEVMSQVQVLLYFLQL 1252


RESULT 14
OPH1_HUMAN
ID   OPH1_HUMAN     STANDARD;      PRT;   802 AA.
AC   O60890; Q8WX47;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Oligophrenin 1.
GN   OPHN1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND DISEASE.
RC   TISSUE=Fetal brain;
RX   MEDLINE=98241228; PubMed=9582072;
RA   Billuart P., Bienvenu T., Ronce N., des Portes V., Vinet M.C.,
RA   Zemni R., Roest Crollius H., Carrie A., Fauchereau F., Cherry M.,
RA   Brillaut S., Hamel B., Fryns J.-P., Beldjord C., Kahn A., Moraine C.,
RA   Chelly J.;
RT   "Oligophrenin-1 encodes a rhoGAP protein involved in X-linked mental
RT   retardation.";
RL   Nature 392:923-926(1998).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANTS THR-45 AND MET-301.
RX   PubMed=10818214;
RA   Billuart P., Chelly J., Carrie A., Vinet M.C., Couvert P.,
RA   McDonell N., Zemni R., Kahn A., Moraine C., Beldjord C., Bienvenu T.;
RT   "Determination of the gene structure of human oligophrenin-1 and
```

```
RT   identification of three novel polymorphisms by screening of DNA from
RT   164 patients with non-specific X-linked mental retardation.";
RL   Ann. Genet. 43:5-9(2000).
RN   [3]
RP   SEQUENCE OF 613-802 FROM N.A.
RA   Wray P.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Stimulates GTP hydrolysis of members of the Rho
CC       family. Could activates GTPase targets that are known to affect
CC       cell migration and outgrowth of axons and dendrites.
CC   -!- TISSUE SPECIFICITY: Expressed in brain.
CC   -!- DISEASE: Defects in OPHN1 are the cause of X-linked non-specific
CC       mental retardation (XLMR) type 60 (MRX60).
CC   -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ001189; CAA04579.1; -.
DR   EMBL; AJ248245; CAB96181.1; -.
DR   EMBL; AJ248246; CAB96181.1; JOINED.
DR   EMBL; AJ248247; CAB96181.1; JOINED.
DR   EMBL; AJ248248; CAB96181.1; JOINED.
DR   EMBL; AJ248249; CAB96181.1; JOINED.
DR   EMBL; AJ248250; CAB96181.1; JOINED.
DR   EMBL; AJ248251; CAB96181.1; JOINED.
DR   EMBL; AJ248252; CAB96181.1; JOINED.
DR   EMBL; AJ248253; CAB96181.1; JOINED.
DR   EMBL; AJ248254; CAB96181.1; JOINED.
DR   EMBL; AJ248255; CAB96181.1; JOINED.
DR   EMBL; AJ248256; CAB96181.1; JOINED.
DR   EMBL; AJ248257; CAB96181.1; JOINED.
DR   EMBL; AJ248258; CAB96181.1; JOINED.
DR   EMBL; AJ248259; CAB96181.1; JOINED.
DR   EMBL; AJ248260; CAB96181.1; JOINED.
DR   EMBL; AJ248261; CAB96181.1; JOINED.
DR   EMBL; AJ248262; CAB96181.1; JOINED.
DR   EMBL; AJ248263; CAB96181.1; JOINED.
DR   EMBL; AJ248264; CAB96181.1; JOINED.
DR   EMBL; AJ248265; CAB96181.1; JOINED.
DR   EMBL; AJ248266; CAB96181.1; JOINED.
DR   EMBL; AJ248267; CAB96181.1; JOINED.
DR   EMBL; AL157700; CAD18899.1; -.
DR   Genew; HGNC:8148; OPHN1.
DR   MIM; 300127; -.
DR   InterPro; IPR001849; PH.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00169; PH; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00233; PH; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS50003; PH_DOMAIN; 1.
KW   GTPase activation; Polymorphism.
FT   DOMAIN      265    368       PH.
FT   DOMAIN      387    534       RHO-GAP.
FT   DOMAIN      629    761       PRO-RICH.
FT   VARIANT      45     45       A -> T.
FT                                /FTId=VAR_013638.
FT   VARIANT     301    301       T -> M.
FT                                /FTId=VAR_013639.
SQ   SEQUENCE   802 AA;  91640 MW;  A24D150048071608 CRC64;

  Query Match             8.0%;  Score 260.5;  DB 1;  Length 802;
  Best Local Similarity   22.2%;  Pred. No. 3.6e-09;
  Matches  164;  Conservative  120;  Mismatches  249;  Indels  207;  Gaps   37;

Qy         12 FEQLVRRV-----EILSEGNEV------------QFIQLAKDFE-DFRKKWQRTDHELG- 52
              :|| : |      ::: :|| :           :| |  : |: ||      || |:
Db         25 YEQELERTNKFIKDVIKDGNALISAMRNYSSAVQKFSQTLQSFQFDFIGD-TLTDDEINI 83

Qy         53 -----KYKDLLMKAETER-----SALDVKLKHARN----QVDVEIKRRQRAEADCEK--- 95
                   :: :|| : | ||     :| |: :|   |    |:    :|::: | | |:
Db         84 AESFKEFAELLNEVENERMMMVHNASDLLIKPLENFRKEQIGFTKERKKKFEKDGERFYS 143

Qy         96 -LERQIQLIREMLMCD-TSGSIQLSEEQ----KSALAFLNRGQPSSSNAGNKRLSTID-- 147
               |:| : |  :          :|: :|:    :|:| :: : |       :|: : ::
Db        144 LLDRHLHLSSKKKESQLQEADLQVDKERHNFFESSLDYVYQIQEVQE---SKKFNIVEPV 200

Qy        148 ----ESGSILSDISFDKTDESLDWDSSLVKTFKLKKREKRRSTSRQFVDGPPGPVKKTRS 203
                   |  | : :: : | : | :   |      |:      |::|: ::     :|:
Db        201 LAFLHSLFISNSLTVELTQDFLPYKQQL--QLSLQNTRNHFSSTREEMEELKKRMKE--- 255

Qy        204 IGSAVDQGNESIVAKTTVTVPNDGGP-IEA--------VSTIETVPYWTRSRRKTGTLQP 254
                         |  |  :|  | | ||           |  | |: :  ::| ||
Db        256 ------------APQTCKLP--GQPTIEGYLYTQEKWALGISWVKYYCQYEKETKTL-- 298

Qy        255 WNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLSL 314
                  |:   : :|       | |    : |   | :   | ||     ||
Db        299 -----TMTPMEQKP-------GAKQGPLDLTLKYCVRR---KTES---IDKRF------- 333

Qy        315 KCRDCRVVSHP---------ECRDRCPLPCI---PTLIGTPVKIGEGMLADFVSQTSPMI 362
                | |       |       |   | : :     :  :|:   : |  : |
Db        334 -CFDIETNERPGTITLQALSEANRRLWMEAMDGKEPIYHSPITKQQEMELNEVGF----- 387

Qy        363 PSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVK---TVPLLSKVDDIHAIC 419
                |  |:| || :|:   |||  | :  |::|   |   |    |   :    ||  |
Db        388 -KFVRKCINIIETKGIKTEGLYRTVGSNIQVQKLLNAFFDPKCPGDVDFHNSDWDIKTIT 446

Qy        420 SLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIH 479
                | || :|||| ||::|:||::  : ||:  : |  : |::  | :||: ||: |  |: |
Db        447 SSLKFYLRNLSEPVMTYRLHKELVSAAKSDNLDYRLGAIHSLVYKLPEKNREMLELLIRH 506

Qy        480 LQRVAQSPHTK---MDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPL 536
                |  | :  |:|   |  :|:  :||||::       | |    :|| |  ||| |    :
Db        507 LVNVCE--HSKENLMTPSNMGVIFGPTLMR---AQEDTVAAMMNIKFQNIVVEIL----I 557

Qy        537 EYWSQFMM--VEQENIDPL-----------------HVIENSNAFSTP-----------Q 566
               |:: :  :   |:    |:                    ::     | |           |
Db        558 EHFGKIYLGPPEESAAPPVPPPRVTARRHKPITISKRLLRERTVFYTSSLDESEDEIQHQ 617

Qy        567 TPDIKV-SLLGPVTTPEHQLL------------KTPSS---------------SSLSQRV 598
               ||:  : | : |   |:|  |            |:||                 |  |:
Db        618 TPNGTITSSIEPPKPPQHPKLPIQRSGETDPGRKSPSRPILDGKLEPCPEVDVGKLVSRL 677

Qy        599 RSTLTKNTPRFGSKSKSATN 618
               :   || ||:       |||
Db        678 QDGGTKITPK-------ATN 690


RESULT 15
RHG6_MOUSE
ID   RHG6_MOUSE     STANDARD;      PRT;   986 AA.
AC   O54834; Q9QZL8;
DT   15-JUL-1999 (Rel. 38, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Rho-GTPase-activating protein 6 (Rho-type GTPase-activating protein
DE   RhoGAPX-1).
GN   ARHGAP6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., REVISIONS TO 600-601, AND ALTERNATIVE SPLICING.
RC   STRAIN=129/Sv;
RX   MEDLINE=20164286; PubMed=10699171;
RA   Prakash S.K., Paylor R., Jenna S., Lamarche-Vane N., Armstrong D.L.,
RA   Xu B., Mancini M.A., Zoghbi H.Y.;
RT   "Functional analysis of ARHGAP6, a novel GTPase-activating protein for
```

```
RT   RhoA.";
RL   Hum. Mol. Genet. 9:477-488(2000).
RN   [2]
RP   SEQUENCE OF 279-659 FROM N.A. (ISOFORM 1).
RC   STRAIN=129/Sv;
RX   MEDLINE=98086484; PubMed=9417914;
RA   Schaefer L., Prakash S.K., Zoghbi H.Y.;
RT   "Cloning and characterization of a novel rho-type GTPase-activating
RT   protein gene (ARHGAP6) from the critical region for microphthalmia
RT   with linear skin defects.";
RL   Genomics 46:268-277(1997).
CC   -!- FUNCTION: GTPASE ACTIVATOR FOR THE RHO-TYPE GTPASES BY CONVERTING
CC       THEM TO AN INACTIVE GDP-BOUND STATE. COULD REGULATE THE
CC       INTERACTIONS OF SIGNALING MOLECULES WITH THE ACTIN CYTOSKELETON.
CC       PROMOTES CONTINUOUS ELONGATION OF CYTOPLASMIC PROCESSES DURING
CC       CELL MOTILITY AND SIMULTANEOUS RETRACTION OF THE CELL BODY
CC       CHANGING THE CELL MORPHOLOGY (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN RETINA AND LUNG.
CC   -!- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF012273; AAC53522.2; -.
DR   EMBL; AF177664; AAD55086.1; -.
DR   HSSP; Q07960; 1RGP.
DR   MGD; MGI:1196332; Arhgap6.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00324; RhoGAP; 1.
KW   Alternative splicing; GTPase activation; SH3-binding.
FT   SITE        343    353       SH3-BINDING.
FT   DOMAIN      412    573       RHO-GAP.
FT   VARSPLIC    638    659       SPDILQTEVSFSMGGRHSSTDS -> TSSVLPAAGQACSQS
FT                                PASDFTP (IN ISOFORM 2).
FT   VARSPLIC    660    986       MISSING (IN ISOFORM 2).
SQ   SEQUENCE   986 AA;  108786 MW;  40800B6457D34CE1 CRC64;

  Query Match             7.8%;  Score 251.5;  DB 1;  Length 986;
  Best Local Similarity   22.3%;  Pred. No. 1.8e-08;
  Matches  127;  Conservative   93;  Mismatches  200;  Indels  149;  Gaps   24;

Qy         66 SALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEEQKSAL 125
              |   |:|:    |   |::| :  |    :|::          ||    | : :: :
Db        205 SGRSVRLRSVPIQSLSELERARLQEVAFYQLQQD---------CDLGCQITIPKDGQKRK 255

Qy        126 AFLNRGQPSSSNAGNKRLSTIDES-GSILSDISFDKTDESLDWDSSLVKTFKLK---KRE 181
                | :   |     ||    | :: |  || :  :             : :|||   :||
Db        256 KSLRKNLDSLGKEKNKDKEFIPQAFGMPLSQVIAND-----------RAYKLKQDLQRE 303

Qy        182 KRRSTSRQFVDG--PPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETV 239
              :::  |  ||    | |  |: : : |:    | |:   :|   ||:     : :| |
Db        304 EQKDASSDFVSSLLPFGNKKQNKELSSS----NSSL--SSTSETPNES---TSPNTPEPA 354

Qy        240 PYWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPES 299
              |   |:||: | :                  ||:          | |  |: :   :
Db        355 P---RARRR-GAM-----------------SVDSI--------TDLDDNQSRLLEALQL 384

Qy        300 CVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTS 359
              :|   : |      | || ::  :|  |
Db        385 SLPAEAQSK----KEKARDKKLSLNPIYR---------------------------- 409

Qy        360 PMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAIC 419
                :| :|  |   :|: ||   |::|:    : |::|:|:| |   |  | :   :| :
Db        410 -QVPRLVDSCCQHLEKHGLQTVGIFRVGSSKKRVRQLREEFDRGVDV-CLEEEHSVHDVA 467

Qy        420 SLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIH 479
              :|||:|||:: :||||  |  ||:    : : :  :  :   :  || | |||  |:
Db        468 ALLKEFLRDMPDPLLTRELYTAFINTL-LLEPEEQLGTLQLLIYLLPPCNCDTLHRLLQF 526

Qy        480 LQRVAQSPH-----------TKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPK- 526
              |  ||:                ||   ||| :||| :: |   : |     |   | :
Db        527 LSIVARHADDNVSKDGQEVTGNKMTSLNLATIFGPNLL-HKQKSSDKEYSVQSSARAEES 585

Qy        527 -----VVERLLSLPLEYWSQFMMVEQENIDPLHVI----ENSNAFSTPQT-PDIKVSLL- 575
                  ||:::               || | | ::    :|    |  :| ||:   ||
Db        586 TAIIAVVQKMI-------------ENYDSLFMVPPDLQNEVLISLLETDPDVVDYLLR 630

Qy        576 ---GPVTTPEHQLLKTPSSSSLSQRVRST 601
                  ::|:  :|:|  | |:  |  ||
Db        631 RKASQSSSPD--ILQTEVSFSMGGRHSST 657
```

Search completed: July 3, 2003, 09:54:41
Job time : 28 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 3, 2003, 09:56:22 ; Search time 55 Seconds
(without alignments)
1321.373 Million cell updates/sec

Title: US-09-881-736-2
Perfect score: 3243
Sequence: 1 MDTMMLNVRNLFEQLVRRVE.....SKSATNLGRQGNFFASPMLK 632

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 440863 seqs, 114992915 residues

Total number of hits satisfying chosen parameters: 440863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3243	100.0	632	10	US-09-881-736-2	Sequence 2, Appli
2	3230	99.6	632	10	US-09-833-790-413	Sequence 413, App
3	2919	90.0	570	10	US-09-833-790-427	Sequence 427, App
4	2734	84.3	628	10	US-09-881-736-4	Sequence 4, Appli
5	541.5	16.7	681	10	US-09-881-736-6	Sequence 6, Appli
6	294.5	9.1	193	10	US-09-802-127-8	Sequence 8, Appli
7	286.5	8.8	1286	9	US-10-153-668-212	Sequence 212, App
8	286.5	8.8	1354	9	US-10-153-668-470	Sequence 470, App
9	270.5	8.3	2548	10	US-09-851-682A-1	Sequence 1, Appli
10	263.5	8.1	555	9	US-09-764-868-878	Sequence 878, App
11	251	7.7	170	9	US-10-080-960-32	Sequence 32, Appl
12	251	7.7	170	10	US-09-802-127-7	Sequence 7, Appli
13	247.5	7.6	291	9	US-09-764-868-898	Sequence 898, App
14	223.5	6.9	152	9	US-10-132-585-5	Sequence 5, Appli
15	218	6.7	103	9	US-10-080-960-34	Sequence 34, Appl
16	213.5	6.6	429	9	US-09-764-868-1029	Sequence 1029, Ap
17	210.5	6.5	803	10	US-09-770-689A-2	Sequence 2, Appli
18	206.5	6.4	257	9	US-09-764-868-1024	Sequence 1024, Ap
19	205.5	6.3	881	10	US-09-816-860A-2	Sequence 2, Appli
20	205	6.3	726	10	US-09-770-689A-4	Sequence 4, Appli
21	204.5	6.3	780	10	US-09-770-689A-5	Sequence 5, Appli
22	200.5	6.2	1702	9	US-09-854-133-434	Sequence 434, App
23	200.5	6.2	1702	9	US-10-144-649A-434	Sequence 434, App
24	200.5	6.2	1702	10	US-09-738-973-434	Sequence 434, App
25	198.5	6.1	254	9	US-09-764-868-1014	Sequence 1014, Ap
26	197.5	6.1	803	9	US-10-059-585-10	Sequence 10, Appl
27	194	6.0	731	9	US-10-201-481-3	Sequence 3, Appli
28	193.5	6.0	643	9	US-10-097-340-16	Sequence 16, Appl
29	192	5.9	390	9	US-10-097-340-21	Sequence 21, Appl
30	192	5.9	464	9	US-10-097-340-14	Sequence 14, Appl
31	192	5.9	718	9	US-10-097-340-19	Sequence 19, Appl
32	192	5.9	751	9	US-10-097-340-18	Sequence 18, Appl
33	189	5.8	663	9	US-10-080-960-14	Sequence 14, Appl
34	187	5.8	2167	10	US-09-801-368-56	Sequence 56, Appl
35	186.5	5.8	341	9	US-09-764-868-1041	Sequence 1041, Ap
36	179.5	5.5	327	9	US-09-764-868-631	Sequence 631, App
37	179.5	5.5	327	9	US-09-955-999-74	Sequence 74, Appl
38	174.5	5.4	547	10	US-09-802-127-2	Sequence 2, Appli
39	164	5.1	332	9	US-09-764-868-1040	Sequence 1040, Ap
40	160.5	4.9	332	9	US-09-764-868-1225	Sequence 1225, Ap
41	159.5	4.9	188	10	US-09-925-300-1017	Sequence 1017, Ap
42	156.5	4.8	288	9	US-09-764-868-1048	Sequence 1048, Ap
43	154.5	4.8	162	9	US-09-764-868-656	Sequence 656, App
44	148.5	4.6	2099	9	US-10-128-714-3290	Sequence 3290, Ap
45	148.5	4.6	2405	9	US-10-128-714-8290	Sequence 8290, Ap

ALIGNMENTS

RESULT 1
US-09-881-736-2
; Sequence 2, Application US/09881736
; Patent No. US20020076785A1
; GENERAL INFORMATION:
; APPLICANT: Glotzer, Michael
; APPLICANT: Jantsch-Plunger, Verena
; APPLICANT: Romano, Alper
; APPLICANT: Mishima, Masanori
; APPLICANT: Kaitna, Susanne
; TITLE OF INVENTION: Cyk-4 polypeptides, DNA molecules encoding them and their use
; TITLE OF INVENTION: screening methods
; FILE REFERENCE: 0652.2260001/EKS/AES
; CURRENT APPLICATION NUMBER: US/09/881,736
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: EP 00 112 880.0
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: EP 01 110 554.1
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/241,231
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: To be determined
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 632
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-881-736-2

Query Match 100.0%; Score 3243; DB 10; Length 632;
Best Local Similarity 100.0%; Pred. No. 2.7e-239;
Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60

Qy         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120
```

```
Qy        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180

Qy        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240

Qy        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300

Qy        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360

Qy        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420

Qy        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480

Qy        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540

Qy        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600

Qy        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              ||||||||||||||||||||||||||||||||
Db        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
```

RESULT 2
US-09-833-790-413
; Sequence 413, Application US/09833790
; Patent No. US20020068288A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Secrist, Heather
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Indirias, Carol Y.
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.512
; CURRENT APPLICATION NUMBER: US/09/833,790
; CURRENT FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 413
; LENGTH: 632
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-790-413

Query Match 99.6%; Score 3230; DB 10; Length 632;
Best Local Similarity 99.7%; Pred. No. 2.6e-238;
Matches 630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDTMMLNVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMK 60

Qy         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEE 120
```

```
Qy        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKR 180
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        121 QKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSHISFDKTDESLDWDSSLVKTFKLKKR 180

Qy        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVP 240

Qy        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300

Qy        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSP 360

Qy        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICS 420

Qy        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHL 480

Qy        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWS 540
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        481 QRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMLQDIKRQPKVVERLLSLPLEYWS 540

Qy        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 QFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRS 600

Qy        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              ||||||||||||||||||||||||||||||||
Db        601 TLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
```

RESULT 3
US-09-833-790-427
; Sequence 427, Application US/09833790
; Patent No. US20020068288A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Secrist, Heather
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Indirias, Carol Y.
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.512
; CURRENT APPLICATION NUMBER: US/09/833,790
; CURRENT FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 427
; LENGTH: 570
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-790-427

Query Match 90.0%; Score 2919; DB 10; Length 570;
Best Local Similarity 99.8%; Pred. No. 1.3e-214;
Matches 569; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         63 TERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEEQK 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEEQK 60

Qy        123 SALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKREK 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKREK 120
```

```
Qy        183 RRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVPYW 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVPYW 180

Qy        243 TRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVP 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVP 240

Qy        303 CGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSPMI 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSPMI 300

Qy        363 PSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICSLL 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 PSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICSLL 360

Qy        423 KDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQR 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQR 420

Qy        483 VAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWSQF 542
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        421 VAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMLQDIKRQPKVVERLLSLPLEYWSQF 480

Qy        543 MMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRSTL 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 MMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVRSTL 540

Qy        603 TKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              ||||||||||||||||||||||||||||||
Db        541 TKNTPRFGSKSKSATNLGRQGNFFASPMLK 570
```

RESULT 4
US-09-881-736-4
; Sequence 4, Application US/09881736
; Patent No. US20020076785A1
; GENERAL INFORMATION:
; APPLICANT: Glotzer, Michael
; APPLICANT: Jantsch-Plunger, Verena
; APPLICANT: Romano, Alper
; APPLICANT: Mishima, Masanori
; APPLICANT: Kaitna, Susanne
; TITLE OF INVENTION: Cyk-4 polypeptides, DNA molecules encoding them and their use in
; TITLE OF INVENTION: screening methods
; FILE REFERENCE: 0652.2260001/EKS/AES
; CURRENT APPLICATION NUMBER: US/09/881,736
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: EP 00 112 880.0
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: EP 01 110 554.1
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/241,231
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: To be determined
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-881-736-4

Query Match 84.3%; Score 2734; DB 10; Length 628;
Best Local Similarity 84.4%; Pred. No. 2e-200;
Matches 534; Conservative 47; Mismatches 46; Indels 6; Gaps 3;

```
Qy          1 MDTMMLNVRNLFEQLVRRVEILSEGNE-VQFIQLAKDFEDFRKKWQRTDHELGKYKDLLM 59
              ||| |:|:  |||||||:||::|||| ::|||: |||||||||:|||: || |:||||:
Db          1 MDTTMVNLWTLFEQLVRRMEIINEGNESIEFIQVVKDFEDFRKKYQRTNQELEKFKDLLL 60

Qy         60 KAETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSE 119
              |||| ||||||||||||||||||||||||||:| |||:||||||::|||||||||||||
Db         61 KAETGRSALDVKLKHARNQVDVEIKRRQRAEAECAKLEQQIQLIRDILMCDTSGSIQLSE 120

Qy        120 EQKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKK 179
              |||||||||||||| || :||| ||||||||||||||||||||||||||||| ||:||
Db        121 EQKSALAFLNRGQASSGHAGNNRLSTIDESGSILSDISFDKTDESLDWDSSLVKNFKMKK 180

Qy        180 REKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETV 239
              |||||| |||||:|||||||||| |||| ||| ||||||||||||:||||||||||||:
Db        181 REKRRSNSRQFIDGPPGPVKKTCSIGSTVDQANESIVAKTTVTVPSDGGPIEAVSTIETL 240

Qy        240 PYWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPES 299
              | ||||| |:| ||| |||| |||| ||||||:||::|||||: ||||||||||||||||
Db        241 PSWTRSRGKSGPLQPVNSDSALNSRPLEPRTDTDNLGTPQNTGGMRLHDFVSKTVIKPES 300

Qy        300 CVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTS 359
              ||||||||||||||||||||:|||||||||||||||| |:|||||||||||||||| |
Db        301 CVPCGKRIKFGKLSLKCRDCRLVSHPECRDRCPLPCIPPLVGTPVKIGEGMLADFVSQAS 360

Qy        360 PMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAIC 419
              ||||:||| |||||||||||| ||||||||||||||||||||:|||||||||||||| ||
Db        361 PMIPAIVVSCVNEIEQRGLTEAGLYRISGCDRTVKELKEKFLKVKTVPLLSKVDDIHVIC 420

Qy        420 SLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIH 479
              |||||||||||||||| |::|||||||||||||| ||||||| ||||||||||||||||
Db        421 SLLKDFLRNLKEPLLTFWLSKAFMEAAEITDEDNSTAAMYQAVSELPQANRDTLAFLMIH 480

Qy        480 LQRVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYW 539
              ||||:||| ||||:||||||||||||| |||||||| |||||| ||||||||||||||
Db        481 LQRVSQSPDTKMDIANLAKVFGPTIVAHTVPNPDPVTMFQDIKRQLKVVERLLSLPLEYW 540

Qy        540 SQFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVR 599
              :|||||:|||||     :  |  |||:|||:|||||||||||| ||:||| |||||||:  .
Db        541 NQFMMVDQENIDS----QRGNGNSTPRTPDVKVSLLGPVTTPEFQLVKTPLSSSLSQRLY 596

Qy        600 STLTKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
              : |:|:|||||:|||||||||:|| || :| ||
Db        597 N-LSKSTPRFGNKSKSATNLGQQGKFFPAPYLK 628
```

RESULT 5
US-09-881-736-6
; Sequence 6, Application US/09881736
; Patent No. US20020076785A1
; GENERAL INFORMATION:
; APPLICANT: Glotzer, Michael
; APPLICANT: Jantsch-Plunger, Verena
; APPLICANT: Romano, Alper
; APPLICANT: Mishima, Masanori
; APPLICANT: Kaitna, Susanne
; TITLE OF INVENTION: Cyk-4 polypeptides, DNA molecules encoding them and their use
; TITLE OF INVENTION: screening methods
; FILE REFERENCE: 0652.2260001/EKS/AES
; CURRENT APPLICATION NUMBER: US/09/881,736
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: EP 00 112 880.0
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: EP 01 110 554.1
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/241,231
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: To be determined
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 681
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-881-736-6

Query Match 16.7%; Score 541.5; DB 10; Length 681;

```
  Best Local Similarity   25.9%;  Pred. No. 6.4e-33;
  Matches  180;  Conservative  109;  Mismatches  285;  Indels  121;  Gaps   22;

Qy          7 NVRNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKW-------QRTDHELGKYKDLLM 59
              | |::|  ::       :  ::    |  : |  || |       :| : :: : :: |
Db         14 NSRHIFNMILNSQRPQFDIKDIGMFHLIDEIERLRKLWKDSEESKKRLNADMREAEEALA 73

Qy         60 KAETERSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSE 119
              ||  : :  |: :|  :  :   ::  :  : |    | : : ::: :      |  |:: 
Db         74 KARKKLAMFDIDVKDTQKHLRALMEENKALKLDLNVYETREKQLKDAMKNGIFNS--LTK 131

Qy        120 EQKSALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWD---SSLVKTFK 176
              | :    ||:          :||:       :: |   |: |  :|:|    |  :
Db        132 EDRDQFKFLHE---PLVRTYSKRVQ--QRHPHLMEDTQDDEDDSEVDYDETGDSFEEVIH 186

Qy        177 LKK-REKRRSTS-----------------------------RQFVDGPPG----PVKK 200
              |:  || |||::                            :|  |      | |:
Db        187 LRNGREVRRSSAAGNAVGGKRRSASAHAITAAANSKRSRSRVMTATIDEEPNEGGTPPKR 246

Qy        201 TRSIGSAVDQGNESIVAKTTVT------VPNDGGP-------------IEAVSTIETVP 240
              |  ||   |   :    || |       |   |                :: : :  |
Db        247 CRDDGSTPHQEMTTTTTTTTTTTIHNSRAQNQDPPRVSLHRQLTRRSLSCGSIPSCDQTP 306

Qy        241 YWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESC 300
              | :    |      : |||: | |:  |   : || : :  || | |:  :     |
Db        307 GQTTNNIGLGMSSAILTKSTLDIRTLKRGTPAWTNGTTR-DIAMRPHTFIEAGIKAMRKC 365

Qy        301 VPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIP----------TLIGTPVKIGEGM 350
              |   :|    |:|||||  | |  | ::  |||||          | |     ||
Db        366 DKCATALKLA-TSMKCRDCHQVVHRSCCNKLHLPCIPRPKTMMTPKSALRGAKPGAGEFR 424

Qy        351 LADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLS 410
              | |  :   ||||: |:|||  :| ||||: |:||: |  |||  | :: || |||| :
Db        425 LQDLCTSAKPMIPAAVIHCVVALEARGLTQEGIYRVPGQVRTVNVLLDE-LRSKTVPNVG 483

Qy        411 KVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEI--TDEDNSIAAMYQAVGELPQA 468
              : |:  |   || |||:|||:||:     :  : || :  || ||   |: : : |||||
Db        484 -LHDVEVITDTLKRFLRDLKDPLIPRTSRQELIVAANLYSTDPDNGRLALNRVICELPQA 542

Qy        469 NRDTLAFLMIHLQRV-AQSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQ-------D 520
              ||||||:| || ::| |||   ||:   :|::  | :: |       ||  ||       |
Db        543 NRDTLAYLFIHWRKVIAQSSRNKMNCEAMARMVAPAVMGH------PVKQSQSQAIAGRD 596

Qy        521 IKRQPKVVERLLSLPLEYWSQFMMVEQENIDPLHVIENSNAFSTPQTPD----IKVSLLG 576
              :  :  |      || :|:           |   ||   | :  |       |:||
Db        597 ATDCHRAMTALFEFDDVYWQRFLGTSA-------VSMASNQIETARHQDNFALCDRSILG 649

Qy        577 PVTTPEHQLLKTPSSSSLSQRVRSTLTKNTPRFGS 611
              ||||       :|::  |::   :|  :     ||
Db        650 PVTT-------SPATPLLARSANATRARGAHLLGS 677


RESULT 6
US-09-802-127-8
; Sequence 8, Application US/09802127
; Patent No. US20020045212A1
; GENERAL INFORMATION:
;  APPLICANT: Glucksmann, Maria Alexandra
;  APPLICANT:  Meyers, Rachel
;  TITLE OF INVENTION: No. US20020045212A1el Human GTPase Activator Proteins
;  FILE REFERENCE: 035800/158994
;  CURRENT APPLICATION NUMBER: US/09/802,127
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: 60/185,611
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 193
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Rho-Gap3 consensus sequence

US-09-802-127-8

  Query Match             9.1%;  Score 294.5;  DB 10;  Length 193;
  Best Local Similarity   40.1%;  Pred. No. 7.5e-15;
  Matches   81;  Conservative   26;  Mismatches   54;  Indels   41;  Gaps    8;

Qy        359 SPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVD----- 413
              || || ||  |:  :|:|||   |:||:||    ||||:| |   :    |  :|
Db          1 SP-IPIIVEKCIEYLEKRGLDTEGIYRVSGSKSRVKELREAFDSGEDD--LDSLDESITE 57

Qy        414 --------DIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEI-------------TDE 451
                      |:| :  ||| :|| | |||||| |   |:|||::              ||
Db         58 ESEDLEEYDVHDVAGLLKLYLRELPEPLLTFELYEEFIEAAKLYQIEATSRKQSEKSEDE 117

Qy        452 DNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQ-SPHTKMDVANLAKVFGPTIVAHAVP 510
              :  : |: : :  || ||| || :| :|| |||: |   ||   ||| |||||::
Db        118 EERLRALRELLSLLPPANRATLRYL-LHLNRVAEHSEVNKMTARNLAIVFGPTLLR---- 172

Qy        511 NPDPVTMSQDIKRQPKVVERLL 532
                |:|   ||| | |||| |:
Db        173 --PPLT---DIKHQNKVVETLI 189


RESULT 7
US-10-153-668-212
; Sequence 212, Application US/10153668
; Publication No. US20030092616A1
; GENERAL INFORMATION:
;  APPLICANT: HONDA, Goichi
;  APPLICANT:  MATSUDA, Akio
;  APPLICANT:  MURAMATSU, Shuji
;  APPLICANT:  ISHIZAWA, Kenya
;  TITLE OF INVENTION: STAT6 Activating Gene
;  FILE REFERENCE: 1254-0207P
;  CURRENT APPLICATION NUMBER: US/10/153,668
;  CURRENT FILING DATE:  2002-05-24
;  PRIOR APPLICATION NUMBER: US 60/293,172
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/316,031
;  PRIOR FILING DATE: 2001-08-31
;  PRIOR APPLICATION NUMBER: US 60/328,403
;  PRIOR FILING DATE: 2001-10-12
;  PRIOR APPLICATION NUMBER: JP 2001-157043
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: JP 2001-260681
;  PRIOR FILING DATE: 2001-08-30
;  PRIOR APPLICATION NUMBER: JP 2001-313175
;  PRIOR FILING DATE: 2001-10-10
;  NUMBER OF SEQ ID NOS: 488
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 212
;   LENGTH: 1286
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-153-668-212

  Query Match             8.8%;  Score 286.5;  DB 9;  Length 1286;
  Best Local Similarity   22.1%;  Pred. No. 4.8e-13;
  Matches  143;  Conservative   98;  Mismatches  216;  Indels  191;  Gaps   26;

Qy         48 DHE-LGKYK-DLLMKAETERSAL----------DVKL--KHARNQVDVEIKRRQRAEADC 93
              ||| :|    |    ::|||| |          | ||  ||  :   ::|   |::  |
Db        187 DHESVGPPSLDAQPNSKTERSKSYDEGLDDYREDAKLSFKHVSSLKGIKIADSQKSSEDS 246

Qy         94 ----------------------------EKLERQIQLIREMLMCDTSGSIQLSEEQKS 123
                                         :::   |:  ::| :     |: | ::::
Db        247 GSRKDSSSEVFSDAAKEGWLHFRPLVTDKGKRVGGSIRPWKQMYVVLRGHSLYLYKDKRE 306

Qy        124 ALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKT---------------------- 161
                  || | ||              | |||: :|
Db        307 QTTPSEEEQPISVNA------------CLIDISYSETKRKNVFRLTTSDCECLFQAEDR 353

Qy        162 DESLDWDSSLVKTFKLKKREKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTV 221
```

```
Db        354 DDMLAWIKTIQESSNLNE-EDTGVTNRDLIS---RRIKEYNNLMSKAEQ-----LPKT-- 402

Qy        222 TVPNDGGPIEAVSTIETVPYWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSN 281
Db        403 -------PRQSLSIRQTL-LGAKSEPKT---QSPHSPKEESERKL---LSKDDTSPPKDK 448

Qy        282 GGMR--LHDFVSKTV-IKPESCVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPT 338
Db        449 GTWRKGIPSIMRKTFEKKPTATGTFGVRL----------DDCP------ 481

Qy        339 LIGTPVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKE 398
Db        482 ----------------PAHTNRYIPLIVDICCKLVEERGLEYTGIYRVPGNNAAISSMQE 525

Qy        399 KFLRVKTVPLLSKVD-------DIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDE 451
Db        526 ELNKG-----MADIDIQDDKWRDLNVISSLLKSFFRKLPEPLFTNDKYADFIEANRKEDP 580

Qy        452 DNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQ-SPHTKMDVANLAKVFGPTIVAHAVP 510
Db        581 LDRLKTLKRLIHDLPEHHYETLKFLSAHLKTVAENSEKNKMEPRNLAIVFGPTLVRTSED 640

Qy        511 NPDPVTMSQDIKRQP---KVVERLLSLPLEYWSQFMMVEQENIDPLHVIENSNAFSTPQT 567
Db        641 N-----MTHMVTHMPDQYKIVETLIQ-----HHDWFFTEEGAEEPLTTVQEESTVDSQPV 690

Qy        568 PDIKVSLLGPVTTPEHQLLKT----PSSSSLSQRVRSTLTKNTPRFGS 611
Db        691 PNI----------DHLLTNIGRTGVSPGDVSDSATSDSTKSKGSWGS 727
```

RESULT 8
US-10-153-668-470
; Sequence 470, Application US/10153668
; Publication No. US20030092616A1
; GENERAL INFORMATION:
; APPLICANT: HONDA, Goichi
; APPLICANT: MATSUDA, Akio
; APPLICANT: MURAMATSU, Shuji
; APPLICANT: ISHIZAWA, Kenya
; TITLE OF INVENTION: STAT6 Activating Gene
; FILE REFERENCE: 1254-0207P
; CURRENT APPLICATION NUMBER: US/10/153,668
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/293,172
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/316,031
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/328,403
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: JP 2001-157043
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: JP 2001-260681
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: JP 2001-313175
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 488
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 470
; LENGTH: 1354
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-668-470

Query Match 8.8%; Score 286.5; DB 9; Length 1354;
Best Local Similarity 22.1%; Pred. No. 5.2e-13;
Matches 143; Conservative 98; Mismatches 216; Indels 191; Gaps 26;

```
Qy         48 DHE-LGKYK-DLLMKAETERSAL----------DVKL--KHARNQVDVEIKRRQRAEADC 93
Db        255 DHESVGPPSLDAQPNSKTERSKSYDEGLDDYREDAKLSFKHVSSLKGIKIADSQKSSEDS 314

Qy         94 ------------------------------EKLERQIQLIREMLMCDTSGSIQLSEEQKS 123
Db        315 GSRKDSSSEVFSDAAKEGWLHFRPLVTDKGKRVGGSIRPWKQMYVVLRGHSLYLYKDKRE 374

Qy        124 ALAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKT---------------------- 161
Db        375 QTTPSEEEQPISVNA------------CLIDISYSETKRKNVFRLTTSDCECLFQAEDR 421

Qy        162 DESLDWDSSLVKTFKLKKREKRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTV 221
Db        422 DDMLAWIKTIQESSNLNE-EDTGVTNRDLIS---RRIKEYNNLMSKAEQ-----LPKT-- 470

Qy        222 TVPNDGGPIEAVSTIETVPYWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSN 281
Db        471 -------PRQSLSIRQTL-LGAKSEPKT---QSPHSPKEESERKL---LSKDDTSPPKDK 516

Qy        282 GGMR--LHDFVSKTV-IKPESCVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPT 338
Db        517 GTWRKGIPSIMRKTFEKKPTATGTFGVRL----------DDCP------ 549

Qy        339 LIGTPVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKE 398
Db        550 ----------------PAHTNRYIPLIVDICCKLVEERGLEYTGIYRVPGNNAAISSMQE 593

Qy        399 KFLRVKTVPLLSKVD-------DIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDE 451
Db        594 ELNKG-----MADIDIQDDKWRDLNVISSLLKSFFRKLPEPLFTNDKYADFIEANRKEDP 648

Qy        452 DNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQ-SPHTKMDVANLAKVFGPTIVAHAVP 510
Db        649 LDRLKTLKRLIHDLPEHHYETLKFLSAHLKTVAENSEKNKMEPRNLAIVFGPTLVRTSED 708

Qy        511 NPDPVTMSQDIKRQP---KVVERLLSLPLEYWSQFMMVEQENIDPLHVIENSNAFSTPQT 567
Db        709 N-----MTHMVTHMPDQYKIVETLIQ-----HHDWFFTEEGAEEPLTTVQEESTVDSQPV 758

Qy        568 PDIKVSLLGPVTTPEHQLLKT----PSSSSLSQRVRSTLTKNTPRFGS 611
Db        759 PNI----------DHLLTNIGRTGVSPGDVSDSATSDSTKSKGSWGS 795
```

RESULT 9
US-09-851-682A-1
; Sequence 1, Application US/09851682A
; Patent No. US20020091248A1
; GENERAL INFORMATION:
; APPLICANT: Adams, Arwen E.
; APPLICANT: Chiu, Choi Ying
; APPLICANT: Duhl, David
; APPLICANT: Gorman, Susan W.
; APPLICANT: Leng, Song
; APPLICANT: Sheffield, Val
; APPLICANT: Welch, Juliet
; TITLE OF INVENTION: MYOSIN IXA AND CYCLIC NUCLEOTIDE GATED
; TITLE OF INVENTION: CHANNEL-15 (CNGC-15) POLYNUCLEOTIDES, POLYPEPTIDES,
; TITLE OF INVENTION: COMPOSITIONS, METHODS, AND USES THEREOF
; FILE REFERENCE: 200130.442
; CURRENT APPLICATION NUMBER: US/09/851,682A
; CURRENT FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: US/09/172,422
; PRIOR FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2548
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-851-682A-1

Query Match 8.3%; Score 270.5; DB 10; Length 2548;
Best Local Similarity 24.4%; Pred. No. 2.2e-11;
Matches 130; Conservative 75; Mismatches 203; Indels 125; Gaps 24;

```
Qy        143 LSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKREKRRSTSRQFVDGPPGPVKKTR 202
Db       1864 LKSMDE--FLLKKVN-DLDNEDSKKDTLVDVVFKKALKEFRQN-----------IFSFY 1908

Qy        203 SIGSAVDQGN-----------ESIVAKTTVTVPNDG---GPIEA-VSTIETVPYWTRSRR 247
Db       1909 SSALAMDDGKSIRYKDLYALFEQILEKTMRLEQRDSLGESPVRVWVNTFKVFLDEYMNEF 1968

Qy        248 KTGTLQPWNSDSTL-----NSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVP 302
Db       1969 KT--------SDCTATKVPKTERKKRRKKETDLV--EEHNG----HIFKATQYSIPTYCEY 2015

Qy        303 CGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGE-------GMLADFV 355
Db       2016 CSSLIWIMDRASVCKLCKYACHKKC-------CLKTTAKCSKKYDPELSSRQFGVELSRL 2068

Qy        356 SQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDD- 414
Db       2069 TSEDRTVPLVVEKLINYIEMHGLYTEGIYRKSGSTNKIKELRQ---GLDTDAESVNLDDY 2125

Qy        415 -IHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTL 473
Db       2126 NIHVIASVFKQWLRDLPNPLMTFELYEEFLRAMGLQERKETIRGVYSVIDQLSRTHLNTL 2185

Qy        474 AFLMIHLQRVA-QSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLL 532
Db       2186 ERLIFHLVRIALQEDTNRMSANALAIVFAPCIL-RCPDTTDPLQSVQDISKTTTCVE--- 2241

Qy        533 SLPLEYWSQFMMVEQEN--------IDPLHVIENSNAFSTPQTPDIKVSLL--------- 575
Db       2242 ---------LIVVEQMNKYKARLKDISSLEFAENK--------AKTRLSLIRRSMGKGRI 2284

Qy        576 ------GPVTTPEHQLLKTPSSSSLSQRVRSTLTKNTPRFGSKSKSATNLGRQ 622
Db       2285 RRGNYPGP-SSP--VVVRLPSVSDVSE---ETLT-------SEAAMETDITEQ 2324
```

RESULT 10
US-09-764-868-878
; Sequence 878, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ32
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 878
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-868-878

Query Match 8.1%; Score 263.5; DB 9; Length 555;
Best Local Similarity 25.8%; Pred. No. 8.1e-12;
Matches 99; Conservative 57; Mismatches 133; Indels 95; Gaps 17;

```
Qy        195 PGP-------VKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVPYWTRSRR 247
Db         52 PGPGAGDFKKFERTSSSGT---------MSSTEELVDPDGGA----------------- 84

Qy        248 KTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKT--------VIKPES 299
Db         85 ---------GASAFEQADLNGMTPELPVAVP--SGPFR-HEGLSKAARTHRLRKLRTPAK 132

Qy        300 CVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPC-IPTLIGTPVKIGEGMLADF--VS 356
Db        133 CRECNSYVYF--QGAECEECCLACHKKCLETLAIQCGHKKLQGRLQLFGQ----DFSHAA 186

Qy        357 QTSP-MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDI 415
Db        187 RSAPDGVPFIVKKCVCEIERRALRTKGIYRVNGVKTRVEKLCQAFENGKELVELSQASP- 245

Qy        416 HAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEIT----------------DEDNSIAA- 457
Db        246 HDISNVLKLYLRQLPEPLISFRLYHELVGLAKDSLKAEAEAKAASRGRQDGSESEAVAVA 305

Qy        458 ----MYQAVGELPQANRDTLAFLMIHLQRVAQ-SPHTKMDVANLAKVFGPTIVAHAVPNP 512
Db        306 LAGRLRELLRDLPPENRASLQYLLRHLRRIVEVEQDNKMTPGNLGIVFGPTLLR---PRP 362

Qy        513 DPVTMS----QDIKRQPKVVERLL 532
Db        363 TEATVSLSSLVDYPHQARVIETLI 386
```

RESULT 11
US-10-080-960-32
; Sequence 32, Application US/10080960
; Publication No. US20020197695A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Glucksmann, Maria
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 80090, 52874,52880,63497, AND 33425
; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES THEREOF
; FILE REFERENCE: 38155-20044.00
; CURRENT APPLICATION NUMBER: US/10/080,960
; CURRENT FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/242,040
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/242,038
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/241,992
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/242,637
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid
US-10-080-960-32

Query Match 7.7%; Score 251; DB 9; Length 170;
Best Local Similarity 39.8%; Pred. No. 1.3e-11;
Matches 64; Conservative 27; Mismatches 52; Indels 18; Gaps 5;

```
Qy        363 PSIVVHCVNEIE------QRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVD--- 413
Db          1 PIIVEKCVEYIEKLYPLAERGLQEEGIYRVSGSASRVKELREAFDKDGAPDSLELSEKEW 60

Qy        414 -DIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAA--EITDEDNSIAAMYQAV-GELPQAN 469
Db         61 FDVHVVAGLLKLYLRELPEPLIPYDLYEEFIRAAKEQIEDPDERLRALKELLSSKLPRAH 120

Qy        470 RDTLAFLMIHLQRVAQ-----SPHTKMDVANLAKVFGPTIV 505
Db        121 YNTLRYLLTHLNRVAEIYIENSAVNKMNARNLAIVFGPTLL 161
```

RESULT 12
US-09-802-127-7
; Sequence 7, Application US/09802127
; Patent No. US20020045212A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: No. US20020045212A1el Human GTPase Activator Proteins
; FILE REFERENCE: 035800/158994

```
;  CURRENT APPLICATION NUMBER: US/09/802,127
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: 60/185,611
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 170
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Rho-Gap consensus sequence
US-09-802-127-7

  Query Match             7.7%;  Score 251;  DB 10;  Length 170;
  Best Local Similarity   39.8%;  Pred. No. 1.3e-11;
  Matches   64;  Conservative   27;  Mismatches   52;  Indels   18;  Gaps    5;

Qy        363 PSIVVHCVNEIE------QRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVD--- 413
Db          1 PIIVEKCVEYIEKLYPLAERGLQEEGIYRVSGSASRVKELREAFDKDGAPDSLELSEKEW 60

Qy        414 -DIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAA--EITDEDNSIAAMYQAV-GELPQAN 469
Db         61 FDVHVVAGLLKLYLRELPEPLIPYDLYEEFIRAAKEQIEDPDERLRALKELLSSKLPRAH 120

Qy        470 RDTLAFLMIHLQRVAQ-----SPHTKMDVANLAKVFGPTIV 505
Db        121 YNTLRYLLTHLNRVAEIYIENSAVNKMNARNLAIVFGPTLL 161


RESULT 13
US-09-764-868-898
; Sequence 898, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ32
;  CURRENT APPLICATION NUMBER: US/09/764,868
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1510
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 898
;   LENGTH: 291
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-764-868-898

  Query Match             7.6%;  Score 247.5;  DB 9;  Length 291;
  Best Local Similarity   30.0%;  Pred. No. 5.3e-11;
  Matches   89;  Conservative   39;  Mismatches  106;  Indels   63;  Gaps   13;

Qy        263 SRQLEPRTETDSVGTPQ-----SNGGMRLHDFVSKTVIKPE--SCVPCGKRIKFGKLSL- 314
Db          1 SRSLDSPTSSPGAGTRQLVKASSTGTESSDDFEER---DPDLGDGLENGLGSPFGKWTLS 57

Qy        315 ---------------KCRDCR--VVSHPECRD---RCPLPCIPTLI----------GTPV 344
Db         58 SAAQTHQLRRLRGPAKCRECEAFMVSGTECEECFLTCHKRCLETLLILCGHRRLPARTPL 117

Qy        345 KIGEGML---ADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFL 401
Db        118 -FGVDFLQLPRDFPEE----VPFVVTKCTAEIEHRALDVQGIYRVSGSRVRVERLCQAFE 172

Qy        402 RVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDED--------- 452
Db        173 NGRALVELSG-NSPHDVSSVLKRFLQELTEPVIPFHLYDAFISLAKTLHADPGDDPGTPS 231

Qy        453 ---NSIAAMYQAVGELPQANRDTLAFLMIHLQRV-AQSPHTKMDVANLAKVFGPTIV 505
Db        232 PSPEVIRSLKTLLVQLPDSNYNTLRHLVAHLFRVAARFMENKMSANNLGIVFGPTLL 288


RESULT 14
US-10-132-585-5
; Sequence 5, Application US/10132585
; Publication No. US20030055234A1
; GENERAL INFORMATION:
;  APPLICANT: Kapeller-Libermann, Rosanna
;  TITLE OF INVENTION: 26030, A HUMAN RHO-GAP FAMILY MEMBER AND
;  TITLE OF INVENTION:  USES THEREFOR
;  FILE REFERENCE: MPI01-101P1RM
;  CURRENT APPLICATION NUMBER: US/10/132,585
;  CURRENT FILING DATE:  2002-04-25
;  PRIOR APPLICATION NUMBER: 60/286,581
;  PRIOR FILING DATE: 2001-04-25
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 152
;   TYPE: PRT
;   ORGANISM: unknown
;   FEATURE:
;   OTHER INFORMATION: ProDom consensus rhoGAP domain
US-10-132-585-5

  Query Match             6.9%;  Score 223.5;  DB 9;  Length 152;
  Best Local Similarity   39.3%;  Pred. No. 1.4e-09;
  Matches   55;  Conservative   25;  Mismatches   39;  Indels   21;  Gaps    6;

Qy        409 LSKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEA--AEITDEDNSIAA---MYQAVG 463
Db         11 MEEYEDVHTVAGLLKQYFRELPEPLLTYELYEEFIEAAKAQVSDEDERMEALEMLKELIK 70

Qy        464 ELPQANRDTLAFLMIHLQRVAQ-SPHTKMDVANLAKVFGPTIVAHAVPNP--------DP 514
Db         71 LLPEANRETLRYLLKHLSRVAQHSEENKMNAQNLAVVFGPTLL-----RPPDNESSLMDA 125

Qy        515 VTMSQDI--KRQPKVVERLL 532
Db        126 VQAMSDVEYQNQTTVVEFLI 145


RESULT 15
US-10-080-960-34
; Sequence 34, Application US/10080960
; Publication No. US20020197695A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  APPLICANT:  Glucksmann, Maria
;  APPLICANT:  Meyers, Rachel
;  TITLE OF INVENTION: 80090, 52874,52880,63497, AND 33425
;  TITLE OF INVENTION:  METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES THEREOF
;  FILE REFERENCE: 38155-20044.00
;  CURRENT APPLICATION NUMBER: US/10/080,960
;  CURRENT FILING DATE:  2001-10-19
;  PRIOR APPLICATION NUMBER: US 60/242,040
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: US 60/242,038
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: US 60/241,992
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: US 60/242,637
;  PRIOR FILING DATE: 2000-10-23
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 34
;   LENGTH: 103
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Consensus amino acid sequence
US-10-080-960-34
```

Query Match 6.7%; Score 218; DB 9; Length 103;
Best Local Similarity 45.1%; Pred. No. 2.1e-09;
Matches 46; Conservative 21; Mismatches 29; Indels 6; Gaps 3;

```
Qy        409 LSKVDDIHAICSLLKDFLRNLKEPLLTFRLNRAFMEA--AEITDEDNSIAA---MYQAVG 463
              : : :|:| :  ||| : | | |||||: |   |:||  |:::|||  : |   : : :
Db          2 MEEYEDVHTVAGLLKQYFRELPEPLLTYELYEEFIEAAKAQVSDEDERMEALEMLKELIK 61

Qy        464 ELPQANRDTLAFLMIHLQRVAQ-SPHTKMDVANLAKVFGPTI 504
               ||:|||:|| :|: || |||| |   ||:  ||| |||||:
Db         62 LLPEANRETLRYLLKHLSRVAQHSEENKMNAQNLAVVFGPTL 103
```

Search completed: July 3, 2003, 10:05:54
Job time : 57 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 3, 2003, 09:50:57 ; Search time 27 Seconds
(without alignments)
688.714 Million cell updates/sec

Title: US-09-881-736-2
Perfect score: 3243
Sequence: 1 MDTMMLNVRNLFEQLVRRVE.....SKSATNLGRQGNFFASPMLK 632

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	301	9.3	1261	3	US-09-080-855-2	Sequence 2, Appli
2	270.5	8.3	2548	4	US-09-172-422-1	Sequence 1, Appli
3	257.5	7.9	165	5	PCT-US93-03076-10	Sequence 10, Appl
4	230	7.1	140	5	PCT-US93-03076-9	Sequence 9, Appli
5	210	6.5	1513	5	PCT-US93-03076-2	Sequence 2, Appli
6	191	5.9	2071	4	US-09-415-522-6	Sequence 6, Appli
7	183.5	5.7	162	5	PCT-US93-03076-8	Sequence 8, Appli
8	161	5.0	1013	4	US-09-415-522-8	Sequence 8, Appli
9	128.5	4.0	878	4	US-09-735-934A-2	Sequence 2, Appli
10	127.5	3.9	1354	3	US-08-685-871-2	Sequence 2, Appli
11	126.5	3.9	724	1	US-07-906-349A-5	Sequence 5, Appli
12	126.5	3.9	724	1	US-08-167-035-2	Sequence 2, Appli
13	126.5	3.9	724	1	US-08-208-887A-2	Sequence 2, Appli
14	126.5	3.9	724	2	US-08-539-005-2	Sequence 2, Appli
15	126.5	3.9	724	4	US-09-280-598-5	Sequence 5, Appli
16	125.5	3.9	722	3	US-08-390-874C-12	Sequence 12, Appl
17	125.5	3.9	722	4	US-09-265-772-12	Sequence 12, Appl
18	118	3.6	720	3	US-08-899-437-6	Sequence 6, Appli
19	118	3.6	720	4	US-09-126-121-6	Sequence 6, Appli
20	116.5	3.6	942	3	US-08-560-005-6	Sequence 6, Appli
21	116.5	3.6	942	4	US-09-418-540-6	Sequence 6, Appli
22	116	3.6	1146	4	US-08-914-999-6	Sequence 6, Appli
23	115.5	3.6	666	4	US-09-134-001C-4739	Sequence 4739, Ap
24	114.5	3.5	1420	2	US-08-540-804-14	Sequence 14, Appl
25	114.5	3.5	1420	2	US-08-218-265-14	Sequence 14, Appl
26	114.5	3.5	1420	3	US-08-521-872-14	Sequence 14, Appl
27	114.5	3.5	1420	4	US-08-590-399-14	Sequence 14, Appl
28	114	3.5	894	4	US-09-735-934A-4	Sequence 4, Appli
29	113.5	3.5	713	3	US-08-899-437-2	Sequence 2, Appli
30	113.5	3.5	713	4	US-09-126-121-2	Sequence 2, Appli
31	112.5	3.5	942	2	US-08-884-681-4	Sequence 4, Appli
32	112.5	3.5	942	4	US-09-258-643-4	Sequence 4, Appli
33	111	3.4	1003	1	US-08-571-758-4	Sequence 4, Appli
34	111	3.4	1003	1	US-08-909-984A-4	Sequence 4, Appli
35	111	3.4	1003	1	US-08-909-983-4	Sequence 4, Appli
36	110	3.4	976	4	US-09-104-324B-4	Sequence 4, Appli
37	107.5	3.3	1312	4	US-09-345-882-29	Sequence 29, Appl
38	107.5	3.3	1471	3	US-08-755-587-188	Sequence 188, App
39	106.5	3.3	1618	1	US-07-853-913-4	Sequence 4, Appli
40	105.5	3.3	588	4	US-08-235-836C-122	Sequence 122, App
41	105.5	3.3	700	1	US-07-720-589-2	Sequence 2, Appli
42	105.5	3.3	700	2	US-08-785-190-2	Sequence 2, Appli
43	105.5	3.3	700	4	US-08-235-836C-66	Sequence 66, Appl
44	105.5	3.3	700	4	US-08-235-836C-74	Sequence 74, Appl
45	105.5	3.3	700	5	PCT-US92-05539-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-080-855-2
; Sequence 2, Application US/09080855A
; Patent No. 6083721
; GENERAL INFORMATION:
; APPLICANT: Saras, Jan
; APPLICANT: Franzn, Petra
; APPLICANT: Aspenstrm, Pontus
; APPLICANT: Hellman, Ulf
; APPLICANT: Gonez, Leonel Jorge
; APPLICANT: Heldin, Carl-Henrik
; TITLE OF INVENTION: PARG, A GTPASE ACTIVATING PROTEIN WHICH INTERACTS WITH PTPL1
; FILE REFERENCE: L0461/7030
; CURRENT APPLICATION NUMBER: US/09/080,855A
; CURRENT FILING DATE: 1998-05-18
; EARLIER APPLICATION NUMBER: 08/805,583
; EARLIER FILING DATE: 1997-02-25
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-080-855-2

Query Match 9.3%; Score 301; DB 3; Length 1261;
Best Local Similarity 22.4%; Pred. No. 4e-19;
Matches 153; Conservative 120; Mismatches 249; Indels 160; Gaps 33;

```
Qy          9 RNLFEQLVRRVEILSEGNEVQFIQLAKDFEDFRKKWQRTDHE-LGKYKDLLMKAETERSA 67
              | | |: :::||   | :|:  :    : |: |   : |  | | : : |: : :   |
Db        372 RRLEEEALQKVE---EADELYKV-CVTNVEERRNDVENTKREILAQLRTLVFQCDLTLKA 427

Qy         68 LDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEEQK---SA 124
              : | | | ::     :  | ::     ||    |   | :    |     :||:|   :
Db        428 VTVNLFHMQHLQAASLADRLQSLCGSAKLYDPGQEYSEFVKATNS-----TEEEKVDGNV 482

Qy        125 LAFLNRGQPSSSNAGNKRLSTI---DESGSILSDISFDKTDESLDWDSSLVKTFKLKKRE 181
                || |||     |     :   | | | |   :  | :     | ::::
Db        483 NKHLNSSQPSGFGPANSLEDVVRLPDSSNKIEEDRCSNSADIT---GPSFIRSWTFGMFS 539

Qy        182 KRRSTSRQFVDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVPY 241
                ||         |  ::||: |     |||       | |
Db        540 DSEST---------GGSSESRSLDS------ESI-------SPGD--------------- 562

Qy        242 WTRSRRKTGTLQPWNSDSTLNSRQLEPRTET--DSVGTPQS---NGGMRLHDFVSKTVIK 296
              : |   :| :    :|   |: |:    :||  :|:|| :    :     | |  : :
Db        563 FHRKLPRTPSSGTMSSADDLDEREPPSPSETGPNSLGTFKKTLMSKAALTHKF--RKLRS 620

Qy        297 PESCVPCGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPC----IP---TLIGTPVKIGEG 349
```

```
Db        621 PTKCRDCEGIVVF--QGVECEECLLVCHRKCLENLVIICGHQKLPGKIHLFG-------- 670

Qy        350 MLADF--VSQTSP-MIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTV 406
Db        671 --AEFTLVAKKEPDGIPFILKICASEIENRALCLQGIYRVCG-----------NKIKTE 716

Qy        407 PLLSKVD------DI-----HAICSLLKDFLRNLKEPLLTFRLNRAFMEAAE----ITDE 451
Db        717 KLCLALENGMHLVDISEFSSHDICDVLKLYLRQLPEPFILFRLYKEFIDLAKEIQHVNEE 776

Qy        452 DNS-------------------IAAMYQAVGELPQANRDTLAFLMIHLQRVA-QSPHTK 490
Db        777 QETKKNSLEDKKWPNMCIEINRILLKSKDLLRQLPASNFNSLHFLIVHLKRVVDHAEENK 836

Qy        491 MDVANLAKVFGPTIVAHAVPNPD--PVTMSQ--DIKRQPKVVERLLSLPLEYWSQFM--- 543
Db        837 MNSKNLGVIFGPSLIR---PRPQTAPITISSLAEYSNQARLVEFLIT-----YSQKIFDG 888

Qy        544 -MVEQENIDPLHVIENS---NAFSTPQTPDIKVSLLGPVTTPEHQLLKTPSSSSLSQRVR 599
Db        889 SLQPQDVMCSIGVVDQGCFPKPLLSPEERDIERSMKSLFFSSKEDIHTSESESKIFERAT 948

Qy        600 STLTKNTPRFGSKSKSATNLGR 621
Db        949 S--------FEESERKQNALGK 962
```

RESULT 2
US-09-172-422-1
; Sequence 1, Application US/09172422A
; Patent No. 6300485
; GENERAL INFORMATION:
; APPLICANT: Adams, Arwen E.
; APPLICANT: Chiu, Choi Ying
; APPLICANT: Duhl, David
; APPLICANT: Gorman, Susan W.
; APPLICANT: Leng, Song
; APPLICANT: Sheffield, Val
; APPLICANT: Welch, Juliet
; TITLE OF INVENTION: MYOSIN IXA AND CYCLIC NUCLEOTIDE GATED
; TITLE OF INVENTION: CHANNEL-15 (CNGC-15) POLYNUCLEOTIDES, POLYPEPTIDES,
; TITLE OF INVENTION: COMPOSITIONS, METHODS, AND USES THEREOF
; FILE REFERENCE: 200130.442
; CURRENT APPLICATION NUMBER: US/09/172,422A
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2548
; TYPE: PRT
; ORGANISM: Homo sapien
S-09-172-422-1

Query Match 8.3%; Score 270.5; DB 4; Length 2548;
Best Local Similarity 24.4%; Pred. No. 1e-15;
Matches 130; Conservative 75; Mismatches 203; Indels 125; Gaps 24;

```
Qy        143 LSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKREKRRSTSRQFVDGPPGPVKKTR 202
Db       1864 LKSMDE--FLLKKVN-DLDNEDSKKDTLVDVVFKKALKEFRQN-----------IFSFY 1908

Qy        203 SIGSAVDQGN----------ESIVAKTTVTVPNDG---GPIEA-VSTIETVPYWTRSRR 247
Db       1909 SSALAMDDGKSIRYKDLYALFEQILEKTMRLEQRDSLGESPVRVWVVNTFKVFLDEYMNEF 1968

Qy        248 KTGTLQPWNSDSTL-----NSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVP 302
Db       1969 KT-------SDCTATKVPKTERKKRRKKETDLV--EEHNG----HIFKATQYSIPTYCEY 2015

Qy        303 CGKRIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGE-------GMLADFV 355
Db       2016 CSSLIWIMDRASVCKLCKYACHKKC-------CLKTTAKCSKKYDPELSSRQFGVELSRL 2068
```

```
Qy        356 SQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDD- 414
Db       2069 TSEDRTVPLVVEKLINYIEMHGLYTEGIYRKSGSTNKIKELRQ---GLDTDAESVNLDDY 2125

Qy        415 -IHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTL 473
Db       2126 NIHVIASVFKQWLRDLPNPLMTFELYEEFLRAMGLQERKETIRGVYSVIDQLSRTHLNTL 2185

Qy        474 AFLMIHLQRVA-QSPHTKMDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLL 532
Db       2186 ERLIFHLVRIALQEDTNRMSANALAIVFAPCIL-RCPDTTDPLQSVQDISKTTTCVE--- 2241

Qy        533 SLPLEYWSQFMMVEQEN--------IDPLHVIENSNAFSTPQTPDIKVSLL--------- 575
Db       2242 ---------LIVVEQMNKYKARLKDISSLEFAENK--------AKTRLSLIRRSMGKGRI 2284

Qy        576 ------GPVTTPEHQLLKTPSSSSLSQRVRSTLTKNTPRFGSKSKSATNLGRQ 622
Db       2285 RRGNYPGP-SSP--VVVRLPSVSDVSE---ETLT-------SEAAMETDITEQ 2324
```

RESULT 3
PCT-US93-03076-10
; Sequence 10, Application PC/TUS9303076
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; TITLE OF INVENTION: GAP-Associated Protein p190 and
; TITLE OF INVENTION: Transduction
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: 2 Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/03076
; FILING DATE: 19930331
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WHI92-03A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US93-03076-10

Query Match 7.9%; Score 257.5; DB 5; Length 165;
Best Local Similarity 37.9%; Pred. No. 1.8e-16;
Matches 64; Conservative 33; Mismatches 63; Indels 9; Gaps 4;

```
Qy        369 CVNEIEQRGLTETGLYRISGCDRTVKELKEKFLR--VKTVPLLSKVDDIHAICSLLKDFL 426
Db          1 CIREIESRGLNSEGLYRVSGFSDLIEDVKMAFDRDGEKADISVNMYEDINIITGALKLYF 60

Qy        427 RNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQS 486
Db         61 RDLPIPLITYDAYPKFIESAKIMDPDEQLETLHEALKLLPPAHCETLRYLMAHLKRV--T 118
```

```
Qy        487 PHTK---MDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLL 532
                  | |   |:  ||  |||||::    |  | :    ||: |  ||| |:
Db        119 LHEKENLMNAENLGIVFGPTLMRS--PELDAMAALNDIRYQRLVVELLI 165


RESULT 4
PCT-US93-03076-9
; Sequence 9, Application PC/TUS9303076
;  GENERAL INFORMATION:
;    APPLICANT:  Whitehead Institute for Biomedical Research
;    TITLE OF INVENTION:  GAP-Associated Protein p190 and
;    TITLE OF INVENTION:  Transduction
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  2 Militia Drive
;      CITY:  Lexington
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/03076
;      FILING DATE:  19930331
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  WHI92-03A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  140 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US93-03076-9

  Query Match             7.1%;  Score 230;  DB 5;  Length 140;
  Best Local Similarity   38.8%;  Pred. No. 5.4e-14;
  Matches   54;  Conservative   24;  Mismatches   59;  Indels    2;  Gaps    2;

Qy        369 CVNEIEQRGLTETGLYRISGCDRTVKELKEKF-LRVKTVPLLSKVDDIHAICSLLKDFLR 427
                || |||:||: | |:||:||    :: ||  | :  | | ::    |::|:   || : |
Db          1 CVEEIERRGMEEVGIYRVSGVATDIQALKAAFDVNNKDVSVMMSEMDVNALAGTLKLYFR 60

Qy        428 NLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQSP 487
                | ||| |      | |   ::|     : |   :  ||:|| |  ||: ||:|||:
Db         61 ELPEPLFTDEFYPNFAEGIALSDPVAKESCMLNLLLSLPEANLLTFLFLLDHLKRVAEKE 120

Qy        488 H-TKMDVANLAKVFGPTIV 505
                || : ||| |||||::
Db        121 AVNKMSLHNLATVFGPTLL 139


RESULT 5
PCT-US93-03076-2
; Sequence 2, Application PC/TUS9303076
;  GENERAL INFORMATION:
;    APPLICANT:  Whitehead Institute for Biomedical Research
;    TITLE OF INVENTION:  GAP-Associated Protein p190 and
;    TITLE OF INVENTION:  Transduction
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  2 Militia Drive
;      CITY:  Lexington
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/03076
;      FILING DATE:  19930331
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  WHI92-03A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1513 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US93-03076-2

  Query Match             6.5%;  Score 210;  DB 5;  Length 1513;
  Best Local Similarity   19.1%;  Pred. No. 2.3e-10;
  Matches  125;  Conservative   99;  Mismatches  224;  Indels  208;  Gaps   27;

Qy         67 ALDVKLKHARNQVDVEIKRRQRAEAD-------------CEKLERQIQLIREML----- 107
                |:||         :| :: | |  | :                |   : :::|
Db        892 AIDV--------LDNDLSREQLTEGEEIAQEIDGRFTSIPCSHPQHKLELFHPFFKDVVE 943

Qy        108 ----------------MCDTSGSIQLSEEQKSALAFLNRGQPSSSNAGNKRLSTIDESGS 151
                              | |:  :  |    | |   |  | | :          | :
Db        944 KKNIIEATHMYDNVAEACSTTEEVFNSPRAGSPLCNSNL-QDSEEDVEPPSYHLFREDAT 1002

Qy        152 ILSDISFDKTDESLDWD-----SSLVKTFKLKKREKRRSTSRQFVDGPPGPVKK------ 200
                |  :| | :  |:: :      | ::  |: |   |             | |||
Db       1003 -LPSLSKDHSKFSMELEGNDGLSFIMSNFESKLNNK-----------VPPPVKPKPPVHF 1050

Qy        201 --TRSIGSAVDQGNESIVAKTTVT---VPNDG-------GPIEAV-------STIETVPY 241
                |: : | :|||:     |:  :   :| ||        |::||        | :||:
Db       1051 EITKDL-SYLDQGHREGQRKSMSSSPWMPQDGFDPSDYAEPMDAVVKPRNEEENIYSVPH 1109

Qy        242 WTRSRRKTGTLQPWN----------SDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVS 291
                  :: |  |::  |          ||| :::  ||   :  :                ||
Db       1110 -DSTQGKIITIRNINKAQSNGSGNGSDSEMDTSSLERGRKVSA---------------VS 1153

Qy        292 KTVIKPESCVPCGK----RIKF--------GKLSLKCRD---------------------- 318
                | |:    |   |:    |  |         | :  |  |
Db       1154 KPVLYRTRCTRLGRFASYRTSFSVGSDDELGPIRKKEEDQASQGYKGDNAVIPYETDEDP 1213

Qy        319 ---------CRVVSHPECRDRCPLPCI------PTLIGTPVKIGEGMLADFVSQTSPMIP 363
                      |    |: :   | | |          | |      |   |:   | ||
Db       1214 RRRNILRSLRRNTKKPKPK---PRPSITKATWESNYFGVP-------LTTVVTPEKP-IP 1262

Qy        364 SIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICSLLK 423
                :  |:  ||  ||:  |:|||:||    :: |: :| :   : |  |   :: :   :|
Db       1263 IFIERCIEYIEATGLSTEGIYRVSGNKSEMESLQRQFDQDHNLDLAEKDFTVNTVAGAMK 1322

Qy        424 DFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRV 483
                |   | :||: : :    :|| :| | :  : |: : : : |: | :   ::: || ||
Db       1323 SFFSELPDPLVPYSMQIDLVEAHKINDREQKLHALKEVLKKFPKENHEVFKYVISHLNRV 1382

Qy        484 AQSPHTK-MDVANLAKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLS--------- 533
                : :     |   ||:  | ||::     : | :| ::  :    ::|  :
Db       1383 SHNNKVNLMTSENLSICFWPTLMRPDFSSMDALTATRSYQ---TIIELFIQQCPFFFYNR 1439

Qy        534 ----------LPLEYWSQFMMVEQENIDPLHVIENSNAFSTPQTPDIKVSLLGPVT 579
                       ||: |           |||::        | :  :   || |::
```

```
Db       1440 PISEPPGAALAPLQPWHPL-----SPSSPLHLL--------PVSHHLPSHLLQPLS 1482


RESULT 6
US-09-415-522-6
; Sequence 6, Application US/09415522A
; Patent No. 6291660
; GENERAL INFORMATION:
;  APPLICANT: Gaffney, Thomas
;  APPLICANT: Wendland, Juergen
;  APPLICANT: Philippsen, Peter
;  TITLE OF INVENTION: No. 6291660el Fungal Genes Required For No. 6291660mal Growth And
;  TITLE OF INVENTION: Development
;  FILE REFERENCE: CGC2046
;  CURRENT APPLICATION NUMBER: US/09/415,522A
;  CURRENT FILING DATE: 1999-10-08
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 2071
;   TYPE: PRT
;   ORGANISM: Ashbya gossypii
US-09-415-522-6

  Query Match             5.9%;  Score 191;  DB 4;  Length 2071;
  Best Local Similarity   22.4%;  Pred. No. 2.4e-08;
  Matches  128; Conservative   71;  Mismatches  203;  Indels  170;  Gaps   25;

Qy         54 YKDLLMKAETERSALDVK--------LKHARNQVDVEIKRRQRAEA---DCEKLERQ-IQ 101
              |:||:   |   |  | |        | |   :  ::  |:   ||   :  ||: | :
Db       1554 YQDLIPNTEQLPSHDDEKSAHQFGSILFHYGTESSIKAFRKASKEAASNEARKLKFQAMG 1613

Qy        102 LIREMLMCDTSGSIQLSEEQKSALAFLNRGQPSSSNAGNKRLSTIDE---SGSIL----S 154
              |  ::|: :    :|  ::::  |              .          |: |     |:|    :
Db       1614 LFNDILVTEVY-KVQRDQKKQEQL------------------TVQEHEAKRSVLIQHPN 1653

Qy        155 DISFDKTDESLDWDSSLVKTFKLKKREKRRSTSRQFVDGPPGPVKKTRSIG----SAVDQ 210
              :|      |:   ||             ||:|     |          |    ||
Db       1654 KVSVSSASSSVSGSSS-------------GSTAR--TSNPAHAAYALNMAGSLSISAARH 1698

Qy        211 GNESIVAKT-------TVTVPNDGGPIEAVSTIETVPYWTRSRRKTGTLQPWNSDSTLNS 263
              |  |: :::        | | |  |       ||      :    ::  | |:   | |   |
Db       1699 GRSSVSSRSSVISNTATATSPASGASPNQTSTSH---HGGMGKKIGGFLRRPFSISGFTS 1755

Qy        264 RQLEPRTETDSVGTPQSNGGM----------RLHDFVSKTVIKP---ESCVP-------- 302
                :  | :  :    |:|| :           : |    ||||    :||:
Db       1756 SSSQYTTTSVVLSGVQANGSISPYELPELTSEIQDTKIVTVIKTFEIKSCIQINNYRQDP 1815

Qy        303 ----CGKRIKFG--KLSLKCRD-------------CRVVSHPECRDRCPLPCIPTLIGT 342
                  | | :     : :|:| |                |   | :   |          : |
Db       1816 DMMHCFKIVMEDGTQHTLQCMDDADMHEWMKAITLSKRYSFHSK---RFKGKTSNKIFGV 1872

Qy        343 PVKIGEGMLADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRISGCDRTVKELKEKFLR 402
              ||:       |   :   :||:|:|  ::||| ||| | ||||: |   ::  ||  |
Db       1873 PVE-------DVCEREGALIPNIIVKLLDEIELRGLDEVGLYRVPGSVGSINALKNAFDD 1925

Qy        403 VKTVPLLSKVDD-----IHAICSLLKDFLRNLKEPLLTFRLNRAFMEAAEITDEDNSIAA 457
              |     ::|     |: |    | :|| | | | |           |  ||  :|
Db       1926 EGAVHNTFTLEDDRWFEINTIAGCFKLYLRELPESLFT----------NEKVDEFVNIMT 1975

Qy        458 MYQ---------------AVGELPQANRDTLAFLMIHLQRVAQS-PHTKMDVANLAKVF- 500
              |:              :  ||  |   |  | :|| || |   : :|| :||| ||
Db       1976 AYKNHEVDLSQFQNGIKTLLSTLPVFNYHILKRLFLHLNRVHQHVENNRMDASNLAIVFS 2035

Qy        501 ------------GPT------IVAHAVPNPD 513
                         |||      :: | : ||:
Db       2036 MSFINQDDLASTMGPTLGLLQMLLQHLIRNPE 2067


RESULT 7
PCT-US93-03076-8
; Sequence 8, Application PC/TUS9303076
;  GENERAL INFORMATION:
;    APPLICANT:  Whitehead Institute for Biomedical Research
;    TITLE OF INVENTION:  GAP-Associated Protein p190 and
;    TITLE OF INVENTION:  Transduction
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  2 Militia Drive
;      CITY:  Lexington
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/03076
;      FILING DATE:  19930331
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  WHI92-03A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  162 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US93-03076-8

  Query Match             5.7%;  Score 183.5;  DB 5;  Length 162;
  Best Local Similarity   25.0%;  Pred. No. 1.7e-09;
  Matches   38;  Conservative   41;  Mismatches   72;  Indels    1;  Gaps    1;

Qy        369 CVNEIEQRGLTETGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICSLLKDFLRN 428
              |:  ||  ||:  |:||:||    :: |: :| :   : |  | : :: :   :| |
Db          1 CIEYIEATGLSTEGIYRVSGNKSEMESLQRQFDQDHNLDLAEKDETVNTVAGAMKSFFSE 60

Qy        429 LKEPLLTFRLNRAFMEAAEITDEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVAQSPH 488
              | :||: : :    :|| :| | :  : |: : : : |: | :   ::: || ||: :
Db         61 LPDPLVPYSMQIDLVEAHKINDREQKLHALKEVLKKFPKENHEVFKYVISHLNRVSHNNK 120

Qy        489 TK-MDVANLAKVFGPTIVAHAVPNPDPVTMSQ 519
                 |   ||:  | ||::     : | :| ::
Db        121 VNLMTSENLSICFWPTLMRPDFSSMDALTATR 152


RESULT 8
US-09-415-522-8
; Sequence 8, Application US/09415522A
; Patent No. 6291660
; GENERAL INFORMATION:
;  APPLICANT: Gaffney, Thomas
;  APPLICANT: Wendland, Juergen
;  APPLICANT: Philippsen, Peter
;  TITLE OF INVENTION: No. 6291660el Fungal Genes Required For No. 6291660mal Growth
;  TITLE OF INVENTION: Development
;  FILE REFERENCE: CGC2046
;  CURRENT APPLICATION NUMBER: US/09/415,522A
;  CURRENT FILING DATE: 1999-10-08
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
;   LENGTH: 1013
;   TYPE: PRT
;   ORGANISM: Ashbya gossypii
US-09-415-522-8
```

```
  Query Match             5.0%;  Score 161;  DB 4;  Length 1013;
  Best Local Similarity   21.6%;  Pred. No. 5.1e-06;
  Matches   83;  Conservative   61;  Mismatches  128;  Indels  112;  Gaps   15;

Qy        178 KKREKRRSTSRQFVDGPPG----PVKKTRSIGSAVDQ------------GNESIVAKTT 220
              |:||:  |      ||| |    |  ::|  |:: :             |   :  |:
Db        648 KEREQWISVLTTLCDGPGGTAAIPSINSKSEASSLFEQTSISDSSYLGPIANLEAMDATS 707

Qy        221 VTVPNDGGPIEAVSTIETVPYWTRSRRKTGTLQPW------------NSDSTLNSRQLE 267
               | |||  |:   |  |        ||:  :  |:             | :::: |:  :
Db        708 PTRPNDPNPVSLTSEEEK----EVKRRRMKSFFPFKKLATTPTPYAAGNDNASIFSQDDD 763

Qy        268 -PRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLSLKCRDC-RVVSHP 325
              |   |:  |  :|   | |    :                 ||      | | :: |||
Db        764 SPVNATNESGISRSLQSMNLQAQYNAV--------------FG---ADLRSCLQLSSHP 805

Qy        326 ECRDRCPLPCIPTLIGTPVKIGEGMLADFVSQTSPMIPSIVVHCVNEI-EQRGLTETGLY 384
                                          |    |||:|   :  : : ||: | |::
Db        806 ----------------------------YQGKYEIPSVVFRTLEFLYKNRGIQEEGIF 835

Qy        385 RISGCDRTVKELKEKFLRVKTVPLLSKVD----------------DIHAICSLLKDFLRN 428
              |:||    :| |:|:| :   | | :  |                |:: :  ||| :||
Db        836 RLSGSSSLIKSLQEQFDKEYDVDLCNYNDKVSVTPGNENQGGLYVDVNTVSGLLKLYLRK 895

Qy        429 LKEPLLTFRLNRAFMEAAEIT----DEDNSIAAMYQAVGELPQANRDTLAFLMIHLQRVA 484
              |  | : |  : |:|:   |     |:   ||  ::|:    :  :: :| :    : :
Db        896 L--PHMIFG-DAAYMDFKRIVERNGDDSKLIALEFRALVNSGRIAKEYVALMYALFELLV 952

Qy        485 Q----SPHTKMDVANLAKVFGPTI 504
              :    | : ||:: ||  || ||:
Db        953 KITENSKYNKMNLRNLCIVFSPTL 976


RESULT 9
US-09-735-934A-2
; Sequence 2, Application US/09735934A
; Patent No. 6372468
; GENERAL INFORMATION:
;  APPLICANT: LI, Jiayin et al
;  TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL000851
;  CURRENT APPLICATION NUMBER: US/09/735,934A
;  CURRENT FILING DATE:  2000-12-14
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 878
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-735-934A-2

  Query Match             4.0%;  Score 128.5;  DB 4;  Length 878;
  Best Local Similarity   23.0%;  Pred. No. 0.0048;
  Matches   88;  Conservative   35;  Mismatches  121;  Indels  139;  Gaps   18;

Qy         14 QLVRRVEILSEGNEVQFIQLAK-DFEDFRKKWQRTDHELGKYKDLLMKAETERSALDVKL 72
              ||||    : ||: |: :  |   ||||    |   | |  :
Db        107 QLVRSSGDIQEGDLVEVVLSASATFEDF----QIRPHALTVHS----------------- 145

Qy         73 KHARNQVDVEIKRRQRAEADCEKL-ERQIQLIREMLMCDTSGSIQLSEEQKSALAFLNRG 131
                          || | |:   |   |:|: | ||  |   |:  :: | :  |
Db        146 --------------YRAPAFCDHCGEMLFGLVRQGLKCDGCG---LNYHKRCAFSIPN-- 186

Qy        132 QPSSSNAGNKRLSTIDESGSILSDISFD-KTDESLDWDSSLVKTFKLKKREKRRSTSRQF 190
                : | |  :|||    | |: |  |    | |||   :           |  |||:
Db        187 --NCSGARKRRLS----STSLASGHSVRLGTSESLPCTA----------EELSRSTTELL 230

Qy        191 VDGPPGPVKKTRSIGSAVDQGNESIVAKTTVTVPNDGGPIEAVSTIETVPYWTRSRRKTG 250
                 ||               :|  : :: |    | |||
Db        231 PRRPP--------------SSSSSSSASSYT----GRPIEL----------------- 253

Qy        251 TLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTVIKPESCVPCGKRIK-F 309
                        | | |:   |                  | |:  :  :|  |  | | :|
Db        254 -------DKMLLSKVKVP----------------HTFLIHSYTRPTVCQACKKLLKGL 288

Qy        310 GKLSLKCRDCRVVSHPECRDRCPLPCI-PTLIGTPVKIGEGMLADF----------VSQT 358
                :  |:|:||:   |  |  | | |:   ||   | : |    ||          |:
Db        289 FRQGLQCKDCKFNCHKRCATRVPNDCLGEALINGDVPMEEA--TDFSEADKSALMDESED 346

Qy        359 SPMIPSI----VVHCVNEIEQRG 377
              | :||      :|   | |  |
Db        347 SGVIPGSHSENALHASEEEEGEG 369


RESULT 10
US-08-685-871-2
; Sequence 2, Application US/08685871
; Patent No. 6013499
;  GENERAL INFORMATION:
;    APPLICANT:  NARUMIYA, Shuh
;    APPLICANT:  IWAMATSU, Akihiro
;    TITLE OF INVENTION:  RHO TARGET PROTEIN KINASE P160
;    NUMBER OF SEQUENCES:  68
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/685,871
;      FILING DATE:  24-JUL-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 8-184102
;      FILING DATE:  25-JUN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-262553
;      FILING DATE:  14-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16887/845
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1354 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-685-871-2

  Query Match             3.9%;  Score 127.5;  DB 3;  Length 1354;
  Best Local Similarity   22.2%;  Pred. No. 0.012;
  Matches   88;  Conservative   52;  Mismatches  191;  Indels   65;  Gaps   16;

Qy          7 NVRNLFEQLVRRVEILSEG--NEVQFIQLAKDFEDFRKKWQRTDHELGKYKDLLMKAETE 64
              |::  ||: :     |    |::  |   |||:  |||    |          | | | |
Db        983 NLKAAFEKNINTERTLKTQAVNKLAEIMNRKDFKIDRKKANTQD---------LRKKEKE 1033

Qy         65 RSALDVKLKHARNQVDVEIKRRQRAEADCEKLERQIQLIREMLMCDTSGSIQLSEEQKSA 124
                | ::|   | : :  : : |:     |  : | ||: |   |     :|:    |
Db       1034 NRKLQLELNQEREKFNQMVVKHQK-----ELNDMQAQLVEE---CAHRNELQMQLASKE- 1084

Qy        125 LAFLNRGQPSSSNAGNKRLSTIDESGSILSDISFDKTDESLDWDSSLVKTFKLKKRE--K 182
```

```
                    |       :|  : :| |: |  | |:|| :|  :| :     :  |   |
Db       1085 ---------SDIEQLRAKLLDLSDSTSVASFPSADETDGNLP-ESRIEGWLSVPNRGNIK 1134

Qy        183 RRSTSRQFVDGPPGPVKKTRSI-----GSAVDQGNESIVA---KTTVTVPNDGGPIEAVS 234
              |    :|:|      |  :: |          :| | |:|    |      |   | :
Db       1135 RYGWKKQYV------VVSSKKILFYNDEQDKEQSNPSMVLDIDKLFHVRPVTQGDVYRAE 1188

Qy        235 TIETVPYWTRSRRKTGTLQPWNSDSTLNSRQLEPRTETDSVGTPQSNGGMRLHDFVSKTV 294
              | | :|       |   :   |         ::||  : :         |    |:|:
Db       1189 T-EEIP-------KIFQILYANEGECRKDVEMEPVQQAEKTNFQNHKG----HEFIPTLY 1236

Qy        295 IKPESCVPCGK---RIKFGKLSLKCRDCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGML 351
                | :|  | |    :     :|:|| | |  | :  |:      |  :   |     ||
Db       1237 HFPANCDACAKPLWHVFKPPPALECRRCHVKCHRDHLDKKEDLICPCKVSYDVTSARDML 1296

Qy        352 ADFVSQTSPMIPSIVVHCVNEIEQRGLTETGLYRIS 387
                ||         | | | :| :     :|  | |
Db       1297 LLACSQDEQ--KKWVTHLVKKIPKN--PPSGFVRAS 1328
```

RESULT 11
US-07-906-349A-5
; Sequence 5, Application US/07906349A
; Patent No. 5434064
; GENERAL INFORMATION:
; APPLICANT: Schlessinger, Joseph
; APPLICANT: Skolnik, Edward Y.
; APPLICANT: Margolis, Benjamin L.
; TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR
; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE KINASES AN
; TITLE OF INVENTION: TARGET PROTEINS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906,349A
; FILING DATE: 30-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/643,237
; FILING DATE: 18-JAN-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 724 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-906-349A-5

Query Match 3.9%; Score 126.5; DB 1; Length 724;
Best Local Similarity 22.2%; Pred. No. 0.0054;
Matches 50; Conservative 50; Mismatches 106; Indels 19; Gaps 9;

```
Qy        325 PECRDRCPLPCIPTLIGTPVKIGEGMLA--DFVSQTSP--MIPSIVVHCVNEIEQRGLTE 380
              |:|   |||  |    |   : :  |   |   | :|  : | :::  |  ||::||
Db         87 PKPRPPRPLPVAPGSSKTEADVEQQALTLPDLAEQFAPPDIAPPLLIKLVEAIEKKGLEC 146

Qy        381 TGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNR 440
              : |||       : ||::  |   |  :  :: |:| :    | :| :|  |::   :
Db        147 STLYRTQS-SSNLAELRQ-LLDCDTPSVDLEMIDVHVLADAFKRYLLDLPNPVIPAAVYS 204

Qy        441 AFME-AAEITDEDNSIAAMYQAV--GELPQANRDTLAFLMIHLQRVAQ-SPHTKMDVANL 496
                :  | |:   :  |  : : :    :|     || :|: |  :::| |   ::   |
Db        205 EMISLAPEVQSSEEYIQLLKKLIRSPSIPHQYWLTLQYLLKHFFKLSQTSSKNLLNARVL 264

Qy        497 AKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWSQ 541
              :::| | :   :  : |      : :   ||:| |:|     |::
Db        265 SEIFSPMLFRFSAASSD------NTENLIKVIEILISTE---WNE 300
```

RESULT 12
US-08-167-035-2
; Sequence 2, Application US/08167035
; Patent No. 5618691
; GENERAL INFORMATION:
; APPLICANT: Schlessinger, Joseph
; APPLICANT: Skolnick, Edward Y.
; APPLICANT: Margolis, Benjamin L.
; TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR
; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
; TITLE OF INVENTION: KINASES AND NOVEL TARGET PROTEINS
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: 10036-2711
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,035
; FILING DATE: 16-DEC-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-062
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 724 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-167-035-2

Query Match 3.9%; Score 126.5; DB 1; Length 724;
Best Local Similarity 22.2%; Pred. No. 0.0054;
Matches 50; Conservative 50; Mismatches 106; Indels 19; Gaps 9;

```
Qy        325 PECRDRCPLPCIPTLIGTPVKIGEGMLA--DFVSQTSP--MIPSIVVHCVNEIEQRGLTE 380
              |: |   |||  |    |   : :  |   |   | :|  : | :::  |  ||::||
Db         87 PKPRPPRPLPVAPGSSKTEADVEQQALTLPDLAEQFAPPDIAPPLLIKLVEAIEKKGLEC 146

Qy        381 TGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNR 440
              : |||       : ||::  |   |  :  :: |:| :    | :| :|  |::   :
Db        147 STLYRTQS-SSNLAELRQ-LLDCDTPSVDLEMIDVHVLADAFKRYLLDLPNPVIPAAVYS 204

Qy        441 AFME-AAEITDEDNSIAAMYQAV--GELPQANRDTLAFLMIHLQRVAQ-SPHTKMDVANL 496
                :  | |:   :  |  : : :    :|     || :|: |  :::| |   ::   |
Db        205 EMISLAPEVQSSEEYIQLLKKLIRSPSIPHQYWLTLQYLLKHFFKLSQTSSKNLLNARVL 264

Qy        497 AKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWSQ 541
              :::| | :   :  : |      : :   ||:| |:|     |::
```

```
Db        265 SEIFSPMLFRFSAASSD------NTENLIKVIEILISTE---WNE 300


RESULT 13
US-08-208-887A-2
; Sequence 2, Application US/08208887A
; Patent No. 5677421
;  GENERAL INFORMATION:
;    APPLICANT:  Schlessinger, Joseph
;    APPLICANT:  Skolnick, Edward Y.
;    APPLICANT:  Margolis, Benjamin L.
;    TITLE OF INVENTION:  NOVEL EXPRESSION CLONING METHOD FOR
;    TITLE OF INVENTION:  IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
;    TITLE OF INVENTION:  KINASES AND NOVEL TARGET PROTEINS
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PENNIE & EDMONDS
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  10036-2711
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/208,887A
;      FILING DATE:  11-MAR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7683-063
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  724 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-208-887A-2

  Query Match             3.9%;  Score 126.5;  DB 1;  Length 724;
  Best Local Similarity   22.2%;  Pred. No. 0.0054;
  Matches   50;  Conservative   50;  Mismatches  106;  Indels   19;  Gaps    9;

Qy        325 PECRDRCPLPCIPTLIGTPVKIGEGMLA--DFVSQTSP--MIPSIVVHCVNEIEQRGLTE 380
              |: |   ||| |    |   : :  |   |   | :|  : | :::  |  ||::||
Db         87 PKPRPPRPLPVAPGSSKTEADVEQQALTLPDLAEQFAPPDIAPPLLIKLVEAIEKKGLEC 146

Qy        381 TGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNR 440
              : |||       : ||::  |   |  :  :: |:| :    | :| :|  |::   :
Db        147 STLYRTQS-SSNLAELRQ-LLDCDTPSVDLEMIDVHVLADAFKRYLLDLPNPVIPAAVYS 204

Qy        441 AFME-AAEITDEDNSIAAMYQAV--GELPQANRDTLAFLMIHLQRVAQ-SPHTKMDVANL 496
              :  | |:   :  |  : : :    :|     || :|: |  :::| |  ::   |
Db        205 EMISLAPEVQSSEEYIQLLKKLIRSPSIPHQYWLTLQYLLKHFFKLSQTSSKNLLNARVL 264

Qy        497 AKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWSQ 541
              :::| | :   :  : |      : :   ||:| |:|     |::
Db        265 SEIFSPMLFRFSAASSD------NTENLIKVIEILISTE---WNE 300


RESULT 14
US-08-539-005-2
; Sequence 2, Application US/08539005
; Patent No. 5858686
;  GENERAL INFORMATION:
;    APPLICANT:  Schlessinger, Joseph
;    APPLICANT:  Skolnick, Edward Y.
;    APPLICANT:  Margolis, Benjamin L.
;    TITLE OF INVENTION:  NOVEL EXPRESSION CLONING METHOD FOR
;    TITLE OF INVENTION:  IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
;    TITLE OF INVENTION:  KINASES AND NOVEL TARGET PROTEINS
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PENNIE & EDMONDS
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  10036-2711
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/539,005
;      FILING DATE:  4-OCT-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/167,035
;      FILING DATE:  16-DEC-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7683-062
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  724 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-539-005-2

  Query Match             3.9%;  Score 126.5;  DB 2;  Length 724;
  Best Local Similarity   22.2%;  Pred. No. 0.0054;
  Matches   50;  Conservative   50;  Mismatches  106;  Indels   19;  Gaps    9;

Qy        325 PECRDRCPLPCIPTLIGTPVKIGEGMLA--DFVSQTSP--MIPSIVVHCVNEIEQRGLTE 380
              |: |   ||| |    |   : :  |   |   | :|  : | :::  |  ||::||
Db         87 PKPRPPRPLPVAPGSSKTEADVEQQALTLPDLAEQFAPPDIAPPLLIKLVEAIEKKGLEC 146

Qy        381 TGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNR 440
              : |||       : ||::  |   |  :  :: |:| :    | :| :|  |::   :
Db        147 STLYRTQS-SSNLAELRQ-LLDCDTPSVDLEMIDVHVLADAFKRYLLDLPNPVIPAAVYS 204

Qy        441 AFME-AAEITDEDNSIAAMYQAV--GELPQANRDTLAFLMIHLQRVAQ-SPHTKMDVANL 496
              :  | |:   :  |  : : :    :|     || :|: |  :::| |  ::   |
Db        205 EMISLAPEVQSSEEYIQLLKKLIRSPSIPHQYWLTLQYLLKHFFKLSQTSSKNLLNARVL 264

Qy        497 AKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWSQ 541
              :::| | :   :  : |      : :   ||:| |:|     |::
Db        265 SEIFSPMLFRFSAASSD------NTENLIKVIEILISTE---WNE 300


RESULT 15
US-09-280-598-5
; Sequence 5, Application US/09280598
; Patent No. 6391584
;  GENERAL INFORMATION:
;    APPLICANT:  Schlessinger, Joseph
;    APPLICANT:  Skolnik, Edward Y.
;    APPLICANT:  Margolis, Benjamin L.
;    APPLICANT:  App, Harold
```

```
;    TITLE OF INVENTION:  A NOVEL EXPRESSION-CLONING METHOD FOR
;    TITLE OF INVENTION:  IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
;    TITLE OF INVENTION:  KINASES AND NOVEL TARGET PROTEINS
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/280,598
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/252,820
;      FILING DATE:  02-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7683-067
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  724 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-09-280-598-5

  Query Match             3.9%;  Score 126.5;  DB 4;  Length 724;
  Best Local Similarity   22.2%;  Pred. No. 0.0054;
  Matches   50;  Conservative   50;  Mismatches  106;  Indels   19;  Gaps    9;

Qy        325 PECRDRCPLPCIPTLIGTPVKIGEGMLA--DFVSQTSP--MIPSIVVHCVNEIEQRGLTE 380
              |: |   |||  |    |   : :  |   |   | :|  : | :::  |  ||::||
Db         87 PKPRPPRPLPVAPGSSKTEADVEQQALTLPDLAEQFAPPDIAPPLLIKLVEAIEKKGLEC 146

Qy        381 TGLYRISGCDRTVKELKEKFLRVKTVPLLSKVDDIHAICSLLKDFLRNLKEPLLTFRLNR 440
              : |||       : ||::  |   |  :  :: |:| :    | :| :|  |::   :
Db        147 STLYRTQS-SSNLAELRQ-LLDCDTPSVDLEMIDVHVLADAFKRYLLDLPNPVIPAAVYS 204

Qy        441 AFME-AAEITDEDNSIAAMYQAV--GELPQANRDTLAFLMIHLQRVAQ-SPHTKMDVANL 496
                :  | |:   :  |  : : :    :|     || :|: |  :::| |    ::   |
Db        205 EMISLAPEVQSSEEYIQLLKKLIRSPSIPHQYWLTLQYLLKHFFKLSQTSSKNLLNARVL 264

Qy        497 AKVFGPTIVAHAVPNPDPVTMSQDIKRQPKVVERLLSLPLEYWSQ 541
              :::| | :   :  : |      : :   ||:| |:|     |::
Db        265 SEIFSPMLFRFSAASSD------NTENLIKVIEILISTE---WNE 300


Search completed: July  3, 2003, 09:56:51
Job time : 31 secs
```